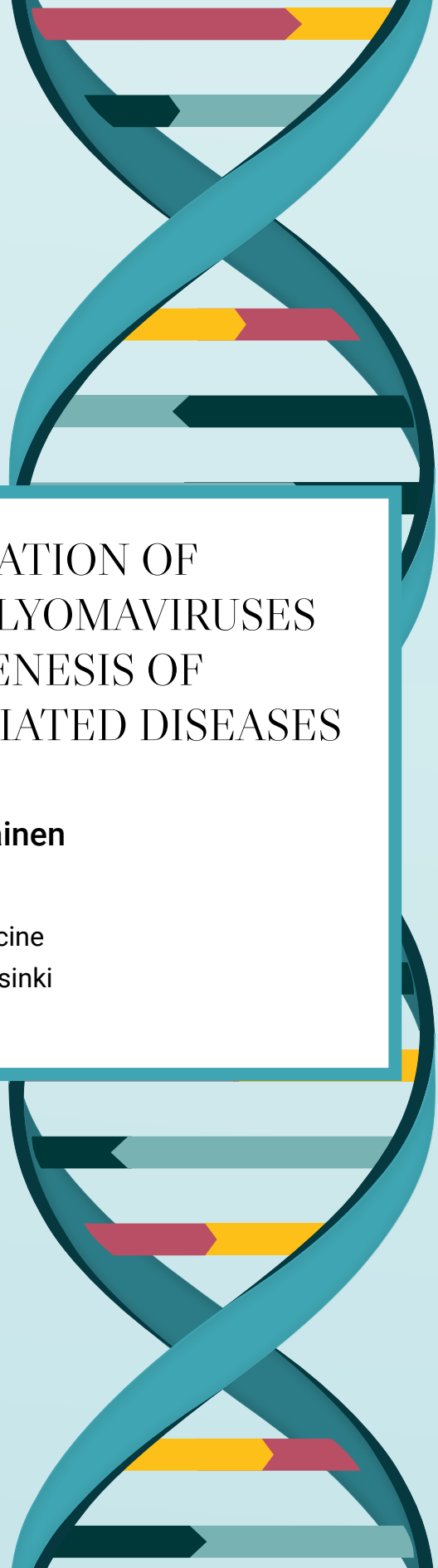


SEQUENCE VARIATION OF HUMAN JC AND BK POLYOMAVIRUSES IN THE PATHOGENESIS OF POLYOMAVIRUS-ASSOCIATED DISEASES

Hanna Liimatainen

Faculty of Medicine
University of Helsinki
2019

SEQUENCE VARIATION OF HUMAN JC AND BK POLYOMAVIRUSES IN THE PATHOGENESIS OF POLYOMAVIRUS-ASSOCIATED DISEASES

HANNA LIIMATAINEN

Department of Virology
Medicum, Faculty of Medicine
University of Helsinki

and

Department of Virology and Immunology
Division of Clinical Microbiology
Helsinki University Hospital Laboratory

ACADEMIC DISSERTATION

To be presented, with the permission of the Faculty of Medicine
of the University of Helsinki, for public examination in Lecture Hall 1,
Haartman Institute, on Friday 10th of January 2020, at 12 noon.

Helsinki 2019

Supervisor

Eeva Auvinen, PhD, Docent
Department of Virology and Immunology
Helsinki University Hospital Laboratory
and
Department of Virology
Medicum, Faculty of Medicine
University of Helsinki
Helsinki, Finland

Reviewers

Varpu Marjomäki, PhD, Docent
Department of Biological and Environmental Science
Nanoscience Center
University of Jyväskylä
Jyväskylä, Finland

Matti Jalasvuori, PhD, Docent
Department of Biological and Environmental Science
Nanoscience Center
University of Jyväskylä
Jyväskylä, Finland

Opponent

Ugo Moens, PhD, Professor
Molecular Inflammation Research Group (MIRG)
Department of Medical Biology
Faculty of Health Sciences
University of Tromsø
Tromsø, Norway

Cover illustration by Niclas Liimatainen

ISBN 978-951-51-5730-0 (paperback)
ISBN 978-951-51-5731-7 (PDF)
Hansaprint
Helsinki 2019

The Faculty of Medicine uses the Urkund system (plagiarism recognition) to examine all doctoral dissertations.

To Niclas and Elias

CONTENTS

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following publications:

I Seppälä H., Virtanen E., Saarela M., Laine P., Paulín L., Mannonen L., Auvinen P., Auvinen E. Single-molecule sequencing revealing the presence of distinct JC polyomavirus populations in patients with progressive multifocal leukoencephalopathy. J Infect Dis 2017, 215(6):889–895

II Seppälä HM., Helanterä IT., Laine PKS., Lautenschlager IT., Paulín LG., Jahnukainen TJ., Auvinen POV., Auvinen E. Archetype JC polyomavirus (JCPyV) prevails in a rare case of JCPyV nephropathy and in stable renal transplant recipients with JCPyV viruria. J Infect Dis 2017, 216(8):981-989

III Virtanen E.*, Seppälä H.*, Helanterä I., Laine P., Lautenschlager I., Paulin L., Mannonen L., Auvinen P., Auvinen E. BK polyomavirus microRNA expression and sequence variation in polyomavirus-associated nephropathy. J Clin Virol 2018, 102:70-76

IV Liimatainen H., Weseslindtner L., Strassl R., Aberle SW., Bond G., Auvinen E. Next-generation sequencing shows marked rearrangements of BK polyomavirus that favor but are not required for polyomavirus-associated nephropathy (Epub ahead of print)

*The authors have contributed equally to the study.

Article III has also been used in the dissertation entitled *Diagnostics and monitoring of chronic virus infections: evaluation of novel molecular methods and markers* by Virtanen E. (Faculty of Medicine, University of Helsinki).

The original publications are referred to in the text by their Roman numerals. The papers I–IV are reprinted with the kind permission of Oxford University Press and Elsevier.

ABBREVIATIONS

ALTO	Alternate frame of the LTag reading frame
BKPVAN	BK polyomavirus-associated nephropathy
BKPyV	BK polyomavirus
bp	Base pair
cART	Combined antiretroviral therapy
CTL	Cytotoxic T lymphocyte
HPyV10	Human polyomavirus 10
HPyV12	Human polyomavirus 12
HPyV6	Human polyomavirus 6
HPyV7	Human polyomavirus 7
HPyV9	Human polyomavirus 9
JCPVAN	JC polyomavirus-associated nephropathy
JCPyV	JC polyomavirus
KIPyV	Karolinska Institute polyomavirus
LIPyV	Lyon IARC polyomavirus
LTag	Large T antigen
MCPyV	Merkel cell polyomavirus
MPS	Massive parallel sequencing
NCCR	Noncoding control region
NGS	Next-generation sequencing
NJPyV	New Jersey polyomavirus
PML	Progressive multifocal leukoencephalopathy
PML-IRIS	PML-associated immune reconstitution inflammatory syndrome
qPCR	Quantitative polymerase chain reaction
SMRT	Single-molecule real-time sequencing
sTag	Small T antigen
STLPyV	Saint Louis polyomavirus
SV40	Simian virus 40
TCR	Transcriptional control region
TSPyV	Trichodysplasia spinulosa polyomavirus
VP1	Viral protein 1
VP2	Viral protein 2
VP3	Viral protein 3
WUPyV	Washington University polyomavirus

ABSTRACT

In this thesis, the role of genetic variation of human JC and BK polyomaviruses (JCPyV, BKPyV) in the development of severe polyomavirus-associated diseases was investigated by characterizing JCPyV and BKPyV strains present in urine, plasma, and cerebrospinal fluid samples using next-generation sequencing. These viruses are established human pathogens causing clinical manifestations almost exclusively in immunosuppressed patients. Immunosuppression allows lytic JCPyV infection in the brain, which is a hallmark of severe brain disease, progressive multifocal leukoencephalopathy (PML). Lytic BKPyV infection in the kidneys may result in the development of BKPyV-associated nephropathy (BKPVAN). Rare cases of nephropathy associated with JCPyV (JCPVAN) have been reported as well. No treatment is currently available leaving reduction of immunosuppression as the only therapy option.

Mutations within the viral genome affect pathogenicity of JCPyV, and the same has been suggested for BKPyV. Although PML is always associated with mutated neurotropic JCPyV strains, the exact content of mutations, as well as where, when, and how neurotropic strains develop is unknown. Mutations are frequently seen in BKPVAN-associated BKPyV strains as well, but their association with the pathogenesis of BKPVAN remains elusive.

First, we performed single-molecule real-time sequencing of complete JCPyV genomes from three PML patients and characterized all mutations within the viral regulatory region and the major capsid protein of each individual virus strain. We identified three distinct neurotropic JCPyV strains from cerebrospinal fluid of one PML patient. Mutations in the viral regulatory region and in the major capsid protein gene seem necessary for PML pathogenesis. Moreover, the presence of an archetype-like strain in one patient indicates that even minor mutations may be sufficient for the development of PML.

Second, we were the first to show the predominance of archetype JCPyV in the urine of twenty stable kidney transplant patients and one patient with extremely rare JCPVAN. Contrary to PML, mutations within the JCPyV genome seem not necessary for the development of JCPVAN, but rather patient-specific factors, such as cell-mediated immunity, may play a role in the pathogenesis of JCPVAN.

In the third and fourth studies, we characterized BKPyV regulatory regions from the plasma and urine samples of kidney transplant patients with or without histologically confirmed BKPVAN. Archetype BKPyV predominated in all but one patient but also very small populations of mutated BKPyV strains were detected. Interestingly, viral microRNA expression, as measured in the plasma of nine presumptive and biopsy-confirmed BKPVAN patients, was increased in the presence of mutated strains. Similar to JCPVAN, archetype strains may be sufficient for the development of BKPVAN, but even low amounts of mutated strains may affect pathogenicity of BKPyV by regulating the expression of viral microRNAs.

This thesis contributes to the understanding of the role of genetic variation of JCPyV and BKPyV in the pathogenesis of PML, JCPVAN, and BKPVAN. Viral mutations seem necessary for the development of PML in the brain, but not for BKPVAN and JCPVAN in the kidney where other factors may play a role. The results help improving the diagnostic methods for the prognosis and prevention of these severe diseases.

TIIVISTELMÄ (ABSTRACT IN FINNISH)

Väitöskirjan tarkoituksena oli tutkia ihmisen JC- ja BK-polyoomavirusten geneettisen muuntelun merkitystä PML-taudissa ja polyoomavirusnefropatiassa sekvensoimalla ja karakterisoimalla viruskantoja potilasnäytteistä. Nämä virukset aiheuttavat vakavia kroonisia tauteja lähes yksinomaan immuunipuutteisilla potilailla, joilla immuunipuolustuksen heikkeneminen mahdollistaa elimistössä piilevien virusten reaktivaation ja lyyttisen virusinfektion. Progressiivinen multifokaalinen leukoenkefalopatia (PML) on seurausta JC-polyoomavirusinfektiosta aivojen gliasoluissa. Lyyttinen BK-polyoomavirusinfektio munuaisissa aiheuttaa BK-polyoomavirusnefropatiaa munuaisensiirtopotilailla. Nefropatia voi harvinaisissa tapauksissa olla myös JC-polyoomavirusinfektion aiheuttama. Polyoomavirus-spesifiä lääkehoitoa ei ole, ja ainoastaan immuunipuolustuksen normalisointi parantaa tautien ennustetta.

Mutaatiot JC-polyoomaviruksen ja mahdollisesti myös BK-polyoomaviruksen perimässä vaikuttavat virusten taudinaiheuttamiskykyyn. PML-tautiin liittyy aina muuntuneen, neurotrooppisen JC-polyoomaviruskannan kehittyminen elimistössä. Ei tiedetä missä, milloin, tai miten neurotrooppiset kannat kehittyvät, tai millaisia mutaatioita vaaditaan PML-taudin kehittymiseen. Mutaatioita tapahtuu myös BK-polyoomaviruksen perimässä, mutta niiden merkitystä BK-polyoomavirusnefropatian kehittymisessä ei tiedetä.

Ensimmäisessä osajulkaisussa karakterisoitiin kaikki JC-polyoomaviruskannat kolmen PML-potilaan selkäydinnesteestä. Työssä osoitettiin, että PML-tautiin voi liittyä samanaikaisesti useita, viruksen geenisäätelyalueen ja kapsidiproteiinigeenin osalta muuntuneita neurotrooppisia kantoja. Toisaalta arkkityyppisen kaltainen viruskanta yhdellä PML-potilaalla viittaisi siihen, että pienetkin viruksen geenisäätelyalueen muutokset olisivat riittäviä PML-taudin kehittymiselle.

Toisessa osajulkaisussa tunnistettiin muuntumattomia, arkkityyppisiä JC-polyoomaviruskantoja kahdenkymmenen oireettoman munuaisensiirtopotilaan ja yhden JC-polyoomavirusnefropatiapotilaan virtsasta. Vaikuttaisi siltä, että toisin kuin PML-taudissa, muutokset viruksen perimässä eivät ole välttämättömiä JC-polyoomavirusnefropatian kehittymiselle vaan muilla tekijöillä, kuten soluvälitteisellä immuniteetilla, on suurempi merkitys.

Kolmannessa ja neljännessä osajulkaisussa osoitettiin, että myös BK-polyoomavirusnefropatian kehittymiseen voisi liittyä arkkityyppinen virus. Kaikkien paitsi yhden tutkitun munuaisensiirtopotilaan plasmassa ja virtsassa arkkityyppiset viruskannat olivat vallitsevia. Useimmilla potilailla havaittiin myös pieniä määriä muuntuneita viruskantoja, ja samoilla potilailla viruksen mikro-RNA-tasot olivat korkeampia. Pienetkin muuntuneiden viruskantojen määrät voisivat siis vaikuttaa BK-viruksen taudinaiheuttamiskykyyn säätelemällä esimerkiksi viruksen mikro-RNA-ekspressiota.

Väitöskirjatyö lisää tietämystä JC- ja BK-polyoomavirusten geneettisen muuntelun merkityksestä PML-taudin ja polyoomavirusnefropatian kehittymisessä: muutokset viruksen perimässä näyttäisivät olevan tärkeitä aivojen PML-taudin, mutta eivät munuaisten polyoomavirusnefropatian kehittymiselle. Saadut tulokset mahdollistavat tehokkaampien diagnostisten menetelmien kehittämisen tautien ennustamiseksi ja ehkäisemiseksi.

INTRODUCTION

Alterations in the immune system function may allow opportunistic viral infections. In the event of prolonged immunosuppression, human polyomaviruses may reactivate from latency and cause severe clinical manifestations. Today, a total of four human polyomaviruses are established human pathogens. The JC polyomavirus (JCPyV) infection in the brain might result in a neurodegenerative brain disease, progressive multifocal leukoencephalopathy (PML) (1). BK polyomavirus (BKPyV), causes BK polyomavirus-associated nephropathy (BKPVAN) of the kidneys (2). Nephropathy may, in rare cases, be caused by the JC polyomavirus (JCPVAN) (3, 4). Merkel cell polyomavirus is currently the only cancer-associated human polyomavirus predisposing to the development of aggressive neuroendocrine cancer, Merkel cell carcinoma (5). Infection of skin cells by Trichodysplasia spinulosa polyomavirus (TSPyV) may lead to the development of rare, non-malignant skin disease (6). As no treatment is currently available, more knowledge on the factors contributing to viral pathogenicity and development of polyomavirus-associated diseases is needed.

Mutations within the viral genome are a prerequisite for PML, and a role of viral mutations have been suggested for BKPVAN as well. Such mutations may alter the pathogenicity of JCPyV: mutations within the noncoding control region (NCCR) may increase JCPyV replication efficiency while mutations in the receptor recognition region of the major capsid protein VP1 may change host cell tropism from kidney epithelial cells to brain glial cells (7-9). BKPyV strains associated with BKPVAN frequently harbor similar mutations that might increase viral replication efficiency (10). Relatively little is known about timing, mechanism, and the exact content of mutations needed for the development of PML or BKPVAN.

Highly sensitive massive parallel sequencing (MPS) techniques enable the detection and characterization of extremely small viral populations present in biological samples. Millions to billions of sequence reads can be produced from single template molecules in hundreds of thousands of parallel reactions in a single run (11). Even very small viral populations and all mutations originating from single viral genomes can be reliably characterized. MPS techniques show promise in molecular diagnostics such as in the detection of drug-resistant HIV strains (12).

In this thesis, sequence variation of JCPyV and BKPyV and its contribution to pathogenesis was studied. For the first time it was shown that multiple mutated JCPyV strains could be associated with PML (I). On the contrary, in the development of JCPVAN and BKPVAN mutations within the JCPyV and BKPyV genomes seem not to be required (II-IV).

1 LITERATURE REVIEW

1.1 Immune system – a sophisticated defense machinery

Immune system has developed to efficiently protect from the constant attack by pathogenic microbes. Two overlapping, yet distinct, defense mechanisms of the immune system exist. Invading pathogens are first encountered by an innate immune system comprising of macrophages, dendritic cells, innate lymphoid cells, neutrophils, eosinophils, basophils, natural killer T cells, various signal molecules such as the complement system, as well as anatomical and physiological barriers such as skin, mucus, and low pH. The cells of innate immunity recognize pathogens by binding to antigens, i.e. pathogen-specific molecules, via so called pattern recognition receptors. Although the innate immune system reacts rapidly to non-self molecules, the response is unspecific and does not lead to the development of immunologic memory against the pathogen.

Some cells of innate immunity, such as dendritic cells, act as antigen-presenting cells and activate an adaptive immune system that consists of T and B lymphocytes. Both lymphocytes are derived from the hematopoietic stem cells of bone marrow; the differentiation of T lymphocytes is finalized in the thymus. Contrary to the innate immunity, the adaptive immunity has a diverse, specific antigen recognition and develops long-lasting immunologic memory against the pathogen. T lymphocytes are responsible for the pathogen-specific, cell-mediated immune response and can be classified into helper, cytotoxic, and regulatory T lymphocytes. Cytotoxic T lymphocytes (CTLs) bind to antigens via their co-receptor CD8 resulting in apoptosis of the infected cell. Activated CD4-positive helper T lymphocytes release cytokines that activate both innate and adaptive immune systems including B lymphocytes. B lymphocytes are responsible for the humoral immune-response and produce pathogen-specific antibodies that bind to antigens and activate cells of innate immunity. Given its specificity, the adaptive immune system plays an important role in controlling the latency of human polyomaviruses and thereby preventing the development of polyomavirus-associated diseases.

1.1.1 Viral infections in immunosuppressed patients

Multiple factors may contribute to the perturbation of the immune system function including acquired viral infections such as HIV, genetics, age, surgery, chemotherapy, and immunosuppressive therapy used to prevent allograft rejection after transplantation or for treatment of disease. Long-lasting immunosuppressive condition may however predispose to opportunistic viral infections caused by latent viruses persisting in the organism. For instance, the clinically most important viral infections after solid-organ transplantation are cytomegalovirus, and Epstein-Barr virus infections (13). Cytomegalovirus is also one of the most important HIV-associated opportunistic viral infections together with hepatitis B and C viruses, and human herpes virus 8, also known as Kaposi's sarcoma-associated herpesvirus (14). Common to all these viruses is the establishment of latency within the organism that can be disturbed by prolonged immunosuppression.

Opportunistic viral infections in immunosuppressed patients have high mortality and morbidity. Reactivation of Epstein-Barr virus may lead to the development of post-transplant lymphoproliferative disease with mortality up to 80% (15). Reactivated cytomegalovirus may cause severe graft-versus-host disease after allogeneic stem cell transplantation (16). Human herpes virus 8 is the etiological agent of Kaposi's sarcoma, an aggressive skin cancer diagnosed in 15-20% of AIDS patients (17). One

major goal during immunosuppressive therapy is thus to prevent and effectively treat such life-threatening viral infections.

1.2 Human polyomaviruses – a group of emerging human pathogens

Human polyomaviruses are restricted human pathogens with potential to cause severe clinical manifestations. Human polyomaviruses belong to the family of Polyomaviridae, a group of DNA tumor viruses infecting mammals, birds, fish, reptiles, amphibians, and invertebrates (18). Originally derived from the Greek words, poly (-multiple) and oma (-tumors), the family includes viruses having potential to cause tumors in their natural host. The research of human polyomaviruses was facilitated by the accidental discovery of a primate polyomavirus, simian virus 40 (SV40), in polio vaccines administered in the early 1960s (19): the finding raised worldwide concern given the oncogenic potential of SV40 in rodents, yet it still remains controversial whether SV40 can indeed be carcinogenic to humans. For decades, JC polyomavirus (JCPyV) and BK polyomavirus (BKPyV) represented the sole human members of the genus, and it was not until the development of the state-of-the-art techniques including rolling-circle amplification and massive parallel sequencing that the discovery of new human polyomaviruses was accelerated. Today, a total of thirteen human polyomaviruses species have been described (20); the putative fourteenth member was discovered in 2017 but has not yet been assigned to a polyomavirus species (21). Although asymptomatic primary infection and the establishment of a latency within a host cell are characteristic to all human polyomaviruses, they are also known to cause several non-malignant neurological, nephrological, and cutaneous diseases. So far, only Merkel cell polyomavirus (MCPyV) has been associated with human cancer (5). The clinical manifestations of human polyomavirus infections are almost exclusively seen after deterioration of immune system function, which is thought to allow the reactivation of human polyomavirus replication.

1.2.1 Epidemiology of human polyomaviruses

Seroprevalence

Human polyomaviruses are ubiquitous; the seroprevalence of most human polyomaviruses exceeds 50% in adults (22) (**Table 1**). The serological data for most human polyomaviruses are inconclusive, however, as serological studies have mainly been conducted on the clinically most relevant human polyomaviruses JCPyV and BKPyV. Approximately 60% of healthy adults have antibodies against JCPyV while the average seroprevalence for BKPyV may reach up to 90% (22-24). Among immunocompetent adults the seroprevalence for both MCPyV and TSPyV is approximately 80% (22-24). The seroprevalences are lowest (~5%) for New Jersey polyomavirus (NJPyV) and the putative novel human polyomavirus, Lyon IARC polyomavirus (LIPyV) (22, 24). High variation in the seroprevalences may indicate geographic differences and different immune reactions against different human polyomaviruses, which in turn suggests variation in the biology of these viruses. Contrary to other human polyomaviruses, the seroprevalence of TSPyV and JCPyV seems to depend on gender-specific factors, as men are more often TSPyV or JCPyV seropositive than women (22-24).

Serological evidence supports the idea that primary infections by most human polyomaviruses occur before the age of ten (23, 25-28). Exceptions include JCPyV, human polyomavirus 7 (HPyV7), human polyomavirus 9 (HPyV9), and NJPyV causing primary infections in the early adolescence or adulthood (23, 26, 27). The varying

onset of primary infection could reflect differences in the biology, transmission, and host susceptibility of different human polyomaviruses. Infection by one human polyomavirus seems not to give immunity to other polyomavirus infections: adults may typically have antibodies against 7 to 9 human polyomaviruses (22, 24).

Due to demographic factors and geographic distribution, the reported seroprevalences of some polyomaviruses, such as Saint Louis polyomavirus (STLPyV), human polyomavirus 10 (HPyV10), and human polyomavirus 12 (HPyV12) may vary considerably (22, 24-26, 28). Different methods including multiplex-PCR and immunofluorescence assays targeting a major capsid protein VP1 (22), and virus-like-particle-based enzyme-linked immunosorbent assays (VLP-ELISA) (24, 25, 27, 28) have been used to detect human polyomavirus-specific antibodies. These methods have drawbacks: most methods are based on VP1 with high amino acid identity among human polyomaviruses, and the consequent cross-reactivity makes it difficult to distinguish between distinct polyomaviruses (29). Human polyomaviruses also share high amino acid similarity in their major regulatory protein, large T antigen (24), hindering its use in seroprevalence studies. Due to the impact of demographical, methodological, and virological factors on the seroprevalence of human polyomaviruses, one can expect more accurate seroprevalences as more serological data accumulates.

Table 1. Seroprevalence and disease associations of thirteen human polyomaviruses and Lyon IARC polyomavirus.

Species name	Name	Year of discovery	Disease association	Seroprevalence in adults (%)
Human polyomavirus 1	BKPyV	1971	BKPVAN, hemorrhagic cystitis	68-99
Human polyomavirus 2	JCPyV	1971	PML, granule cell neuronopathy, meningitis, JCPVAN	25-68
Human polyomavirus 3	KIPyV	2007	Respiratory disease	42-98
Human polyomavirus 4	WUPyV	2007	Respiratory disease	57-99
Human polyomavirus 5	MCPyV	2008	Merkel cell carcinoma	20-86
Human polyomavirus 6	HPyV6	2010	Pruritic and dyskeratotic dermatosis	62-98
Human polyomavirus 7	HPyV7	2010	Pruritic and dyskeratotic dermatosis	36-86
Human polyomavirus 8	TSPyV	2010	Trichodysplasia spinulosa	75-91
Human polyomavirus 9	HPyV9	2011	Unknown	17-70
Human polyomavirus 10	HPyV10	2012	Gastroenteritis	32-100
Human polyomavirus 11	STLPyV	2013	Unknown	57-93
Human polyomavirus 12	HPyV12	2013	Unknown	2-97
Human polyomavirus 13	NJPyV	2014	Unknown	2-56
Human polyomavirus 14	LIPyV	2017	Unknown	5-7

Abbreviations: BKPyV, BK polyomavirus; JCPyV, JC polyomavirus; KIPyV, Karolinska Institute polyomavirus; WUPyV, Washington University polyomavirus; MCPyV, Merkel cell polyomavirus; HPyV6, human polyomavirus 6; HPyV7, human polyomavirus 7; TSPyV, Trichodysplasia spinulosa polyomavirus; HPyV9, human polyomavirus 9; HPyV10, human polyomavirus 10; STLPyV, Saint Louis polyomavirus; HPyV12, human polyomavirus 12; NJPyV, New Jersey polyomavirus; LIPyV, Lyon IARC polyomavirus; BKPVAN; BK polyomavirus-associated nephropathy; PML, progressive multifocal leukoencephalopathy; JCPVAN; JC polyomavirus-associated nephropathy.

Host cells and suggested routes of transmission

Partly due to the high host-receptor variability, the transmission routes of the different human polyomaviruses vary (**Figure 1**). Primary infection of human polyomaviruses is assumed asymptomatic and several transmission routes have been proposed. It seems plausible that transmission takes place in the respiratory or the gastrointestinal tract, or in direct skin contact. First, BKPyV and less frequently JCPyV DNA has been detected in tonsils of healthy children (30, 31) and BKPyV DNA also in nasopharyngeal samples of children both with and without respiratory illness (32). The airborne transmission has also been proposed for Karolinska Institute polyomavirus (KIPyV) and Washington University polyomavirus (WUPyV) detected in the broncho-alveolar lavage and nasopharyngeal aspirates of children; these sites can also serve as sites of primary infection for these viruses (33, 34). Similarly, both HPyV10, human polyomavirus 6 (HPyV6), HPyV7, and MCPyV have been detected in the upper respiratory tract of both children and adults suggesting airborne transmission (35).

Fecal-oral transmission has been suggested for some human polyomaviruses. Replication of HPyV10 may occur in the gastrointestinal tract as it has been detected in the stool of healthy and immunocompromised children (35). The gastrointestinal tract has also been suggested as a site of primary infection for HPyV12 and STLPyV (36, 37). Fecal-oral transmission has been proposed for JCPyV as well (32). In the case of HPyV6, HPyV7, and TSPyV the reactivation of viral replication rather than primary infection could take place in the gastrointestinal tract (35).

Some human polyomaviruses, such as MCPyV, TSPyV, HPyV6, HPyV7, and HPyV9, may spread via direct skin contact as they seem to be part of the skin microbiome in both children and adults (38). NJPyV has been only detected in endothelial and muscle tissues (39). Limited data exist for the most recently discovered human polyomaviruses and larger studies are needed to determine the route of transmission and the site of primary infection.

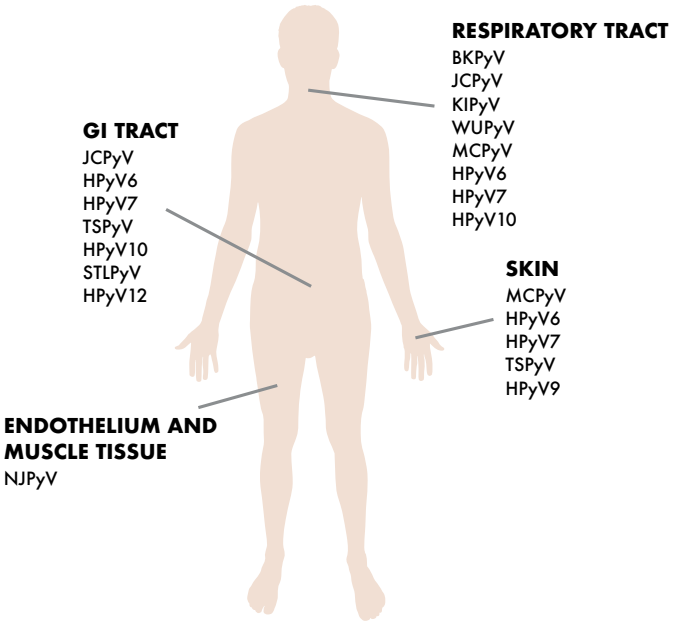


Figure 1. Discovery of human polyomaviruses.

1.2.2 Molecular biology of human polyomaviruses

The structural knowledge of polyomavirus capsid structures comes from SV40 studies. All polyomaviruses have a non-enveloped icosahedral capsid with the diameter of approximately 40-50 nm (40). The capsid encloses a single copy of an enclosed, circular, double-stranded DNA genome molecule. The genome is small, approximately 5,000 base pairs (bp) in size, but it ranges from 4,776 bp (SLPyV) to 5,387 bp (MCPyV) (5, 37). Human polyomaviruses have a highly conserved protein coding region and based on the amino acid identity of the large T antigen (LTag) they can be classified into three genera: alpha-, beta-, and deltapolyomaviruses (**Figure 2**) (18). Within each genus, the classification into different species is based on the difference within the coding sequence of LTag which has to be >15% (41). The genome of all human polyomaviruses has similar organization and it consists of a viral regulatory region, or the noncoding control region (NCCR), as well as early and late viral gene regions activated either before or after the onset of viral genome replication, respectively.

The early and late gene regions encode for regulatory or structural proteins in opposite strands, respectively. All human polyomaviruses express two regulatory proteins, small T antigen (sTag) and LTag, with important roles in the regulation of viral replication and gene expression. Some human polyomaviruses harbor open reading frames for accessory regulatory proteins, middle T antigen (TSPyV, STLPyV) and/or alternate frame of the LTag reading frame, ALTO in short (MCPyV, NJPyV, TSPyV) (37, 42-44). The role of additional regulatory proteins in the viral life cycle is yet to be elucidated but they might regulate host cell cycle rather than viral replication (42). MCPyV and TSPyV also produce truncated spliced variants of LTag, 57kT and 21kT, respectively (43, 45); of note, MCPyV 57kT regulates cellular growth but seems not contribute to transformation.

Besides regulatory proteins, all human polyomaviruses encode for the structural proteins viral protein (VP) 1 and 2, and all except MCPyV encode also VP3 (46). Minor capsid proteins VP2 and VP3 are produced from the same mRNA using different internal in-frame start codon (40). VP1 is a major capsid protein and relatively well conserved but does contain stretches of highly variable regions responsible for the variation seen in the antigenic properties of human polyomaviruses. Further, variation in the host cell receptor recognition region of VP1 may contribute to receptor specificity among human polyomaviruses (47-49).

Bidirectional viral replication and regulation of gene expression by the NCCR is common to all human polyomaviruses. The NCCR has raised scientific and clinical interest due to the high inter- and intra-species variation among human polyomaviruses that may have a role in the viral pathogenicity (50).

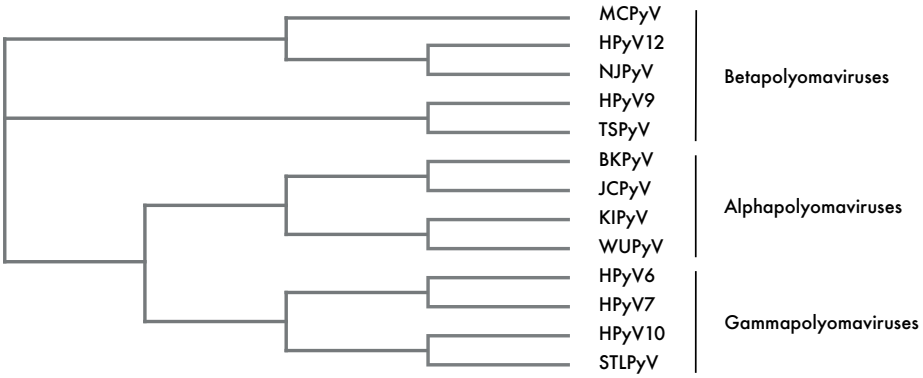


Figure 2. Phylogeny of human polyomaviruses. Clustal Omega multiple sequence alignment tool and Simple Phylogeny tree builder (EMBL-EBI) were used to construct the phylogenic tree based on the nucleic acid identity of LTag among human polyomaviruses.

1.2.3 Polyomaviruses as human pathogens

Following primary infection, all human polyomaviruses establish a latency within an organism. The immune system seems to be crucial in controlling the viral latency, as human polyomaviruses cause severe clinical manifestations almost exclusively in immunosuppressed individuals. To date, four human polyomaviruses (JCPyV, BKPyV, MCPyV, and TSPyV) have been established as etiological agents in human diseases (**Table 1**). JCPyV and BKPyV cause PML and BKPVAN, respectively (1, 2). MCPyV is associated with the development of an aggressive skin cancer, Merkel cell carcinoma (5) and TSPyV infection in the skin cells may lead to the development of a benign but devastating skin disease, trichodysplasia spinulosa (6).

Trichodysplasia spinulosa is a benign hyperproliferative skin disease characterized by alopecia of the eyebrows, and the presence of follicular papules and keratin spines on face and ears (6). A total of 35 trichodysplasia spinulosa cases have been diagnosed in solid organ transplant patients and in patients suffering from hematological malignancies, such as leukemia and lymphoma (51). The pathogenicity of TSPyV has been related to the viral ALTO and middle T antigen proteins (43). Intriguingly, trichodysplasia spinulosa may also result from a primary TSPyV infection (52), different from other human polyomaviruses where reactivation from latency is thought to be required for disease onset.

MCPyV is so far the only human polyomavirus known to be causative in human cancer. MCPyV contributes to the development of majority (80%) of Merkel cell carcinoma cases (5). Merkel cell carcinoma is a rare (incidence ~1/100,000 persons) but highly aggressive neuroendocrine skin cancer with high (86%) recurrence rate (53, 54). The 5-year-survival rate is thus low, approximately 40% (54). Intriguingly, MCPyV-positive Merkel cell carcinomas have better prognosis than MCPyV-negative Merkel cell carcinomas (55). Integration of MCPyV into the host chromosome and truncation of the carboxy-terminal end of LTag, and possibly also in ALTO, seem to enable tumor survival (5, 42, 56). Mutations in the LTag prevent viral DNA replication but do not affect the viral gene expression (57).

The growing body of evidence suggests that other human polyomaviruses may also be pathogenic in humans. HPyV6 or HPyV7 infection may result in pruritic and dyskeratotic dermatosis in immunosuppressed as well as in immunocompetent individuals (58). The causal role of KIPyV or WUPyV in respiratory disease, or HPyV10

in gastroenteritis of immunosuppressed and immunocompetent children has been proposed (59, 60). Whether these human polyomaviruses are indeed human pathogens remains to be validated.

Sequence variation of recently discovered human polyomaviruses

Regardless of the stability and low mutation rate of DNA viruses, mutations frequently take place within human polyomavirus genome, and they may affect viral pathogenicity (**Table 2**). Deletions and insertions occur within the NCCR of Merkel cell carcinoma-associated MCPyV (5), and of possibly dermatosis-associated HPyV6 and HPyV7 (61), respiratory disease-associated KIPyV and WUPyV (59), and gastroenteritis-associated HPyV10 (35). Whether sequence variation takes place in TSPyV strains associated with trichodysplasia spinulosa is unknown. The biological impact of these modifications is to be further investigated but variation in the length and structure of the NCCR among human polyomaviruses seems to confer differences in cell tropism, viral replication and gene expression, and possibly pathogenicity (50).

Human polyomaviruses may also acquire mutations in the viral protein coding region. The dermatosis-associated HPyV6 and HPyV7 strains may have sometimes amino acid point mutations located both in the receptor recognition region of VP1 and functional domains of LTag (61). Occasionally point mutations or larger deletions occur in the VP1, LTag, and ALTO regions of pathogenic MCPyV strains (5, 42, 57, 62). Also KIPyV and WUPyV strains associated with respiratory disease and gastroenteritis-associated HPyV10 strains harbor VP1 and LTag point mutations (35, 59). Apart from MCPyV (56, 57), the pathogenic role of protein coding region mutations in viral pathogenicity remains to be validated.

Table 2. Genetic variation of recently discovered human polyomaviruses associated with human diseases.

Human polyomavirus	Disease association	Sequence variation	Effect on viral pathogenicity
KIPyV	Respiratory disease	NCCR, VP1, VP2, sTag, LTag	Unknown
WUPyV	Respiratory disease	NCCR, VP1, VP2, LTag	Unknown
MCPyV	Merkel cell carcinoma	VP1, LTag, ALTO	Dysregulation of host oncogenes, inhibition of viral replication
HPyV6	Pruritic and dyskeratotic dermatosis	NCCR, VP1, LTag	Unknown
HPyV7	Pruritic and dyskeratotic dermatosis	NCCR, VP1, VP2, VP3, LTag	Unknown
TSPyV	Trichodysplasia spinulosa	Unknown	Unknown
HPyV10	Gastroenteritis	NCCR, VP1, LTag	Unknown

Abbreviations: KIPyV, Karolinska Institute polyomavirus; WUPyV, Washington University polyomavirus; MCPyV, Merkel cell polyomavirus; HPyV6, human polyomavirus 6; HPyV7, human polyomavirus 7; TSPyV, Trichodysplasia spinulosa polyomavirus; HPyV10, human polyomavirus 10; NCCR, noncoding control region; VP1, viral protein 1; VP2, viral protein 2; VP3, viral protein 3; sTag, small T antigen; LTag, large T antigen; ALTO, alternate frame of the LTag reading frame.

The impact of sequence variation in viral pathogenicity has been by far most studied in JCPyV and BKPyV. Mutations within the JCPyV genome are a prerequisite for the development of PML (63, 64) and a role has been proposed for BKPyV variation in BKPVAN. The exact characteristics and content of mutations needed for the pathogenicity, however, remains still elusive. It is also unknown whether individual viral strains harboring a specific set of mutations, similar to HIV drug resistance (12), contribute to the pathogenicity.

1.3 JC and BK polyomaviruses – the first human polyomaviruses to be discovered

JCPyV (former nomenclature JCV) was isolated from the brain tissue of a PML patient suffering from non-Hodgkin lymphoma and was later established as an etiological agent of PML (1). PML is a rapidly progressing, neurodegenerative brain disease with high mortality. The onset of PML is always preceded by the emergence of neurotropic strains having mutations within the JCPyV genome, especially rearrangements within the NCCR (63). These mutations are thought to occur within an individual after asymptomatic primary infection by so called archetype JCPyV with stable NCCR structure (65). NCCR mutations may increase replication efficiency of neurotropic strains in the brain cells and may therefore be necessary for PML pathogenesis (8). Despite of decades of intensive studies, however, the steps from primary infection to PML onset remain largely unsolved.

At the same time, BKPyV (previously known as BKV) was discovered from the urine of kidney transplant patient with ureteric stenosis (2). A subsequent report described BKPyV also in kidney transplant patients suffering from tubulo-intestinal nephritis (66). Today, BKPyV is established as the etiological agent of BKPVAN, a major cause of graft rejection after kidney transplantation. Although BKPyV strains harboring mutations within the NCCR have been detected in association with BKPVAN (10, 67), no clear association with BKPVAN pathogenesis has yet been established.

1.3.1 Molecular biology

Genome structure

The genomes of JCPyV and BKPyV share a high (~70%) nucleic acid identity and their average size is 5,000 bp. In naturally occurring viral variants the genome size may vary which is mainly due to variation within the NCCR: for instance, PML-associated JCPyV Mad-1 and GS/B strains have genome sizes of 5,130 and 5,172 bp, respectively. Similarly, the genomes of BKPVAN-associated BKPyV Gardner and MM strains are 5,190 and 4,963 bp in size, respectively. The genome of BKPyV is packed around cellular histones H2A, H2B, H3, and H4 forming a nucleosome structure (68); whether the same is true for JCPyV is unknown but likely.

JCPyV and BKPyV genomes are structurally similar to other human polyomaviruses. The bidirectional NCCR divides the genome into the early (~2,400 bp) and late viral gene regions (~2,200 bp) (**Figure 3**). The NCCR contains the origin of replication, TATA or TATA-like promoters, enhancer elements, and binding sites for multiple host cell transcription factors that regulate transcription of viral early and late genes and replication of the viral genome (67, 69). As expression of transcription factors varies from tissue to tissue and cell type to cell type, the host cell permissiveness for JCPyV or BKPyV replication depends on the binding sites present in the NCCR of each virus (70, 71). JCPyV has additional regulatory elements

affecting viral replication and gene expression in the viral late gene region (72). The NCCRs of JCPyV and BKPyV enjoy great scientific interest as mutations within the NCCR affect the number of transcription factor binding sites and may thus change viral pathogenicity.

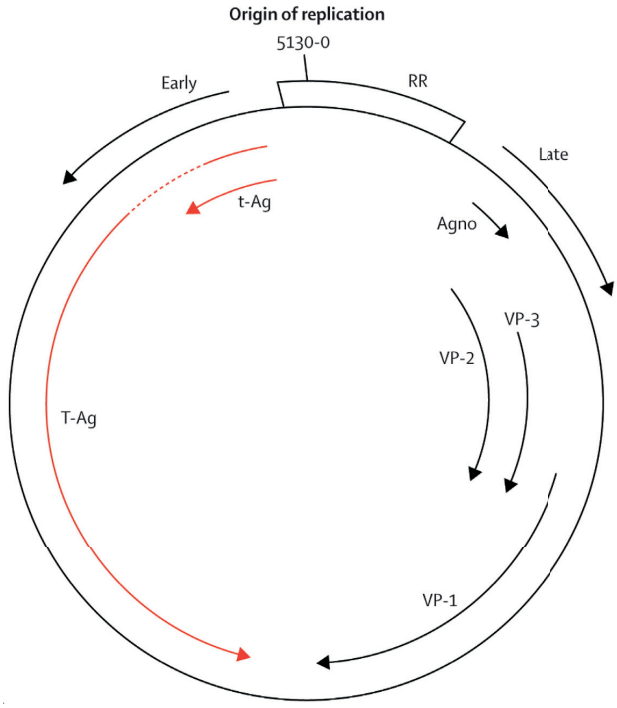


Figure 3. Genome of the neurotropic JCPyV Mad-1 strain. Genome of the prototype BKPyV Gardner strain has similar genomic organization but is 5,190 bp in size. Reproduced with permission from Cinque et al. 2009 (73).

Regulatory proteins

JCPyV and BKPyV encode for LTag and sTag that are produced from the same polycistronic mRNA by alternative splicing. As a major regulatory protein, LTag controls the replication of the viral genome and transcription of viral genes including itself. It has multiple functional domains: J domain, LXCXE (Leu-X-Cys-X-Glu) domain, nuclear localization sequence, and domains for DNA binding, Zinc-binding, and ATPase activity (**Figure 4**) (74). The J domain interacts with cellular heat shock proteins to aid in viral protein synthesis (75) while the LXCXE domain binds to tumor suppressor proteins pRb, p107, p130, and p53 and promotes transition to S phase of the cell cycle (76, 77). To initiate the viral genome replication, LTag binds and unfolds viral DNA thus enabling the binding of the host cell replication machinery to the origin of replication (78, 79). To initiate transcription of viral late genes, LTag binds to the late promoter via TATA-binding protein and associated factors as well as host cell transcription factors such as pur-α, YB-1, and Sp1 (80-82).

Due to the identical amino terminal ends of T antigens, sTag of both JCPyV and BKPyV shares the same J domain with LTag but has a unique hypervariable domain and Zinc-binding domain in the carboxyl terminal end (83). sTag promotes

the initiation and maintenance of S phase by regulating mTOR pathway via its Zinc-binding domain (84, 85). Similar to LTag, sTag also regulates the expression of tumor suppressors p130 and p107 highlighting the importance of cell cycle control in polyomavirus replication (84).

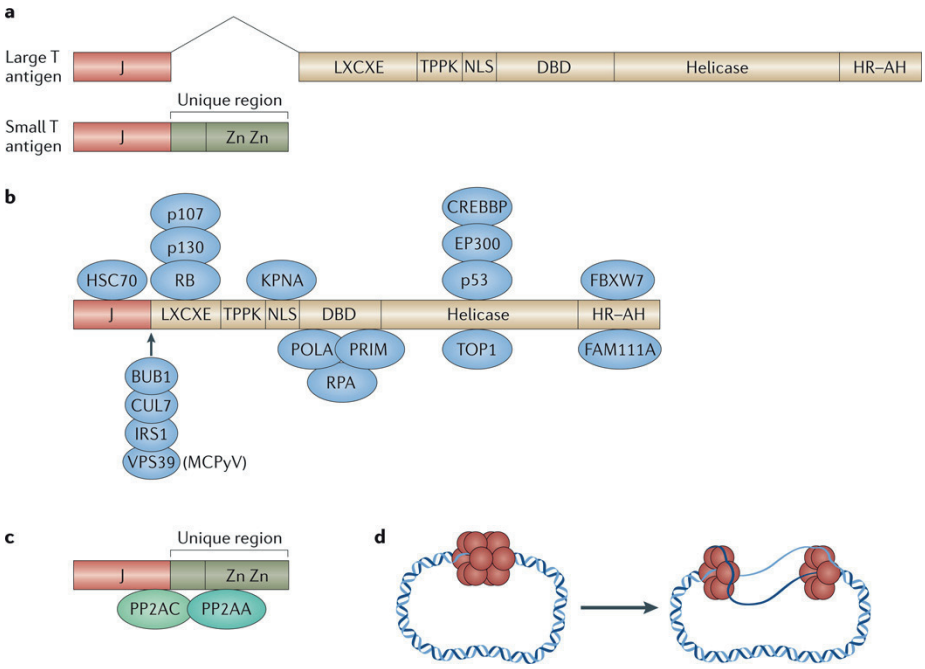


Figure 4. Structure and function of large and small T antigens. The J domain is shared by LTag and sTag but both proteins have unique regions: LXCXE domain, nuclear localization sequence (NLS), DNA-binding domain (DBD), and helicase domain for LTag and Zinc-binding domain for sTag (**a**). Both T antigens bind to multiple host cell factors and regulate e.g. the cell cycle (**b, c**). LTag initiates bidirectional viral DNA replication by binding and unwinding viral DNA at the origin of replication (**d**). Reproduced with permission from DeCaprio and Garcea 2013 (86).

Alternatively-spliced variants of LTag have been identified for both BKPyV and JCPyV. JCPyV produces $T`_{165}$, $T`_{135}$, and $T`_{136}$ –proteins that share J and LXCXE domains with LTag (87). T` proteins play a role in JCPyV replication in regulating e.g. the expression of tumor suppressor proteins p107, p130, and pRb (77). BKPyV has been shown to produce a truncated version of LTag containing J and LXCXE domains of LTag indicating that truncated LTag also regulates viral replication (88).

Unique to JCPyV and BKPyV, both viruses produce cytoplasmic agnoprotein, another auxiliary protein. Agnoprotein is produced at later stages of infection where it inhibits viral DNA replication and increases the expression of viral late genes (89, 90). Both JCPyV and BKPyV agnoproteins act as viroporins by destabilizing the plasma membrane and thereby facilitating the release of progeny virions from the host cell (91, 92).

Capsid proteins

The protein capsid of JCPyV and BKPyV has skewed lattice (T=7d symmetry) and an average diameter of 45 nm (93, 94). The capsid is composed of a total of 72 pentameric capsomers each composed of 5 copies of major capsid protein VP1. Each capsomer is bound via amino-terminal end of VP1 to one copy of minor capsid proteins VP2 and VP3 attached to the viral genome (95).

VP1 harbors amino-terminal nuclear localization and genome encapsidation sequences, antiparallel β-barrel structure with multiple protruding surface loops, and a C-terminal extension (95). Importantly, the surface loops BC, DE, and HI form the receptor recognition region that binds to glycoproteins at the host cell membrane and initiate viral entry into the host cell (96).

Minor capsid proteins VP2 and VP3 bind to viral genome via DNA-binding domain and enhance simultaneous replication and encapsidation of viral genome (97). Transportation of VP1 monomers into nucleus for capsid formation is dependent on the minor capsid proteins (98, 99) and at least BKPyV VP2 seems to have a role in the attachment and entry into the host cell (99).

Viral microRNAs

Similar to many other viruses, JCPyV and BKPyV both encode two noncoding regulatory microRNA molecules: jcv-miR-J1-5p and jcv-miR-J1-3p and bkv-miR-B1-5p and bkv-miR-B1-3p, respectively (100). The 3p microRNAs of JCPyV and BKPyV are identical in their sequence (101). The expression of microRNAs may be conserved among pathogenic human polyomaviruses as microRNAs are also produced by MCPyV (102), although TSPyV microRNAs have not been identified as yet. JCPyV- and BKPyV-encoded microRNAs autoregulate viral early gene expression by binding to a complementary region in the 3´ end of LTag and sTag mRNA and suppressing their translation (100). The expression of BKPyV microRNAs themselves is regulated by the late promoter within the NCCR, implying that mutations within the NCCR might inhibit the normal function of BKPyV microRNAs and subsequently enhance viral replication (103).

JCPyV- and BKPyV-encoded microRNAs have also other roles in the viral life cycle. The 3p microRNAs help the viruses to evade host immune response by inhibiting the expression of stress-induced ligand ULBP3 and preventing the activation of CTLs (101). Viral microRNAs may have a role in the viral latency as JCPyV and BKPyV microRNAs maintain low-level viral replication that would remain unnoticed by the immune system (104, 105). MicroRNAs of both viruses might also serve as viral signal molecules as they are actively transported between infected and uninfected cells (106).

A role for JCPyV and BKPyV microRNAs in viral pathogenicity and development of polyomavirus-associated diseases has been proposed. High-level expression of JCPyV microRNAs, especially jcv-miR-J1-5p, has been detected in the brain of PML patients and in plasma, urine, and cerebrospinal fluid of immunosuppressed patients at risk for PML (100, 107). Similarly, BKPyV microRNAs might be associated with the development of BKPVAN as they have been detected in urine and plasma of BKPVAN patients (108, 109). JCPyV and BKPyV microRNAs may indeed show potential as biomarkers for JCPyV- and BKPyV-associated diseases.

Viral proteins and auxiliary regulatory molecules of JCPyV and BKPyV as well as their functions are summarized in **Table 3**.

Table 3. Proteins and auxiliary regulatory molecules produced by JCPyV and BKPyV.

Protein	Gene region	Size (aa)		Function
		JCPyV	BKPyV	
Large T antigen	Early	688	695	Viral replication and gene expression, cell cycle control
Small T antigen	Early	172	172	Cell cycle control
T'_{135} T'_{136} T'_{165}	Early	135, 136, 165	-	Viral replication, cell cycle control
Truncated large T antigen	Early	-	133	Viral replication, cell cycle control
Agnoprotein	Late	71	66	Viral replication and gene expression, virion egress
VP1	Late	354	362	Major capsid protein
VP2	Late	344	351	Minor capsid protein
VP3	Late	225	232	Minor capsid protein
jcv-miR-J1-5p, bkv-miR-B1-5p	Late	20-22 nt	20-22 nt	Suppression of viral early gene expression
jcv-miR-J1-3p, bkv-miR-B1-3p	Late	20-22 nt	20-22 nt	Suppression of viral early gene expression, immune evasion

Abbreviations: VP1, viral protein 1; VP2, viral protein 2; VP3, viral protein 3; nt, nucleotide; aa, amino acid.

Genotypes

At present, the complete protein coding region of JCPyV is used for genotyping. Seven JCPyV genotypes (1-4, 6-8) as well as subtypes of genotypes 1, 2, and 3 have been distinguished (110). Genotype 4 is probably a recombinant of genotypes 1 and 3 while genotype 5 was shown to be a subtype of genotype 3 rather than a novel genotype (111, 112). Genotypes 1, 2, and 4 are most prevalent in USA (112, 113) while genotypes 3 and 6 are more often identified in individuals of African origin (**Figure 5**) (111, 114). Europeans are commonly infected with JCPyV genotypes 1 and 4, and genotypes 2, 7, and 8 are most frequently detected in Asia (114, 115). Although increased PML risk has not been associated with any specific genotype, some strains of genotype 2 seem to associate with PML more often (113).

Based on the differences within VP1 nucleotides 1744 to 1812, BKPyV isolates are classified into four genotypes (I-IV) (116). Genotypes I and IV can be further divided into multiple subtypes (116, 117). Genotype I is detected in people from Africa, USA, Asia, and Europe while genotype IV is an Asian and European genotype (117-119). Rarer genotypes II and III seem to have a quite specific distribution: genotype II is mainly found in Europe and Africa, and genotype III in Asia and Africa (119). The pathogenicity of the different BKPyV genotypes remains largely unknown although it has been suggested that kidney transplant patients infected with BKPyV genotype I have increased risk to develop BKPVAN (120).

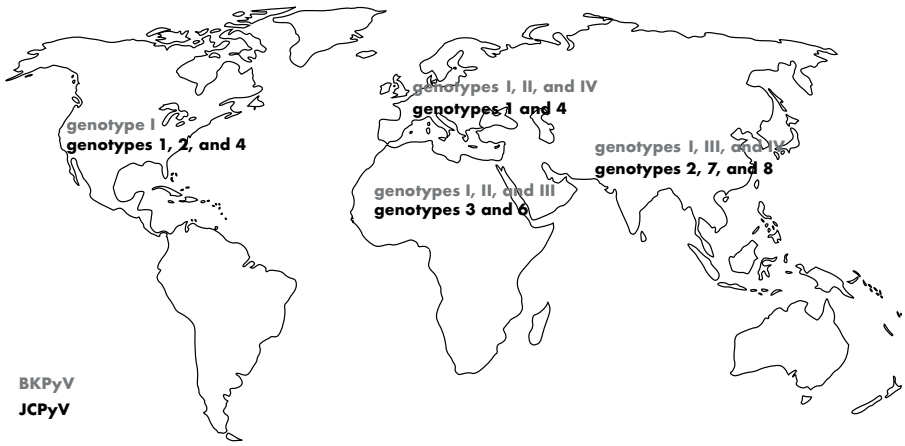


Figure 5. Distribution of JCPyV and BKPyV genotypes.

1.3.2 Viral life cycle

The narrow host cell range of JCPyV and BKPyV places restrictions on studies of their biology due to the lack of animal models. Primary infection is thought to be asymptomatic without apparent clinical symptoms. Viral life cycle starts with an attachment to the viral receptor at the host cell membrane (**Figure 6**). Primary infection is thought to occur in the tonsils and tonsillar lymphocytes (32, 121) from where JCPyV and BKPyV infect leukocytes in order to spread into the systemic circulation (122). JCPyV and BKPyV bind to sialylated glycoproteins to enter the host cell. Upon entry, the receptor recognition region of JCPyV VP1 attaches to the N-linked glycoproteins, oligosaccharide lactoseries tetrasaccharide c, and α2,6-linked sialic acid of GT1b gangliosides (47, 123). The serotonin receptor $5HT_{2A}$ serves as a co-receptor for the glial cell entry (124). BKPyV uses GT1b and GD1b gangliosides, and α2,8-linked sialic acid of N-linked glycoprotein as primary receptors (125, 126); the possible co-receptors are currently unknown.

Following attachment, the virion is internalized and subsequently transported into the nucleus for viral DNA replication. Different strategies are used to enter the host cell: JCPyV utilizes clathrin-dependent endocytosis while BKPyV enters via caveolin-dependent endocytosis (127, 128). After endosomal escape, the virion is transported via microtubules and microfilaments into the endoplasmic reticulum (127, 129). In the endoplasmic reticulum the capsid is partially degraded before transportation into the nucleus with help of the minor capsid proteins (98, 99).

To replicate the viral genome, several cellular signaling pathways are regulated by LTag and sTag to induce S phase of the cell cycle (76, 77, 84, 85). LTag binds to the origin of replication to initiate replication and viral gene expression with the help of host transcription factors and DNA polymerase II (78, 79). At later stages of infection, the cytoplasmic capsid proteins are transported to the nucleus where capsid assembly and genome encapsidation take place (94, 97, 98, 129). Finally, JCPyV and BKPyV are lytic viruses meaning that progeny virions are released from the host cell via e.g. apoptosis (91, 92, 129) and transmitted to the uninfected cells.

After primary infection, JCPyV and BKPyV establish latency within the organism. The factors contributing to the latency remain elusive but epigenetic regulation, such as methylation of viral DNA, may play a role (130, 131). JCPyV and BKPyV remain inactive or slowly replicating in the kidneys, bone marrow, and tonsils (31, 132). Even

the brain has been suggested as a site of latency (133). High seroprevalence but low incidence of PML and BKPVAN in the population indicates that many undiscovered factors play a role in the pathogenicity of JCPyV and BKPyV.

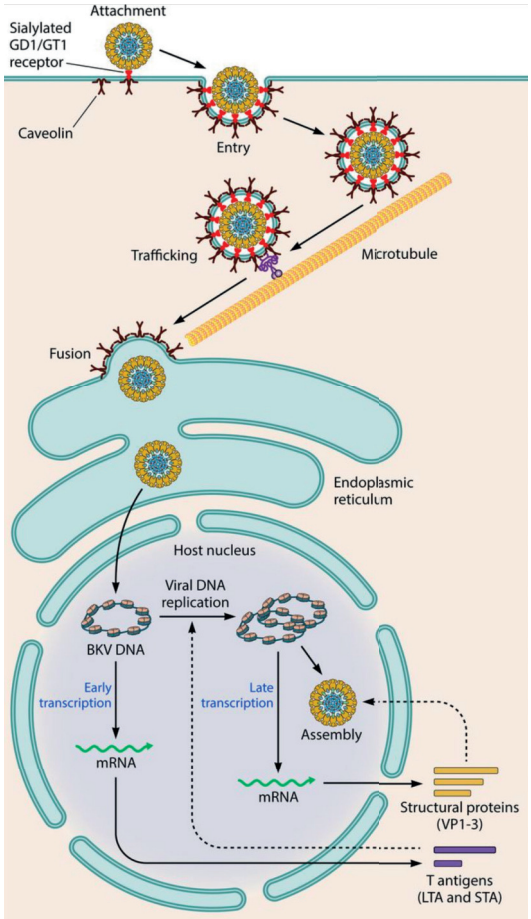


Figure 6. The infection cycle of BKPyV (BKV). The infection cycle of JCPyV is similar but JCPyV uses sialylated GT1 gangliosides, oligosaccharide lactoseries tetrasaccharide c, and N-linked glycoproteins as host cell receptors. The internalization of JCPyV is mediated by clathrin-dependent endocytosis while caveolin-dependent endocytosis is used by BKPyV. The capsid is partially degraded in the endoplasmic reticulum before transportation to the nucleus for viral DNA replication and gene expression. Reproduced with permission from Ambalathingal et al. 2017 (134).

1.4 Progressive multifocal leukoencephalopathy

1.4.1 PML is due to infection of neurotropic JC polyomavirus

PML (Progressive Multifocal Leukoencephalopathy) is a rare neurodegenerative brain disease characterized by progressive damage (-pathy) of white matter (leuko-) of the brain (encephalo-) at multiple locations (multifocal). JCPyV is an etiological agent of PML: demyelination of the white brain matter is due to lytic infection of

glial cells, astrocytes and oligodendrocytes, by neurotropic JCPyV (1). PML has a life-threatening clinical course and, despite huge efforts, no JCPyV-specific antiviral treatment currently exists leaving the modulation of immunosuppression as the only therapy option. This can, however, lead to other complications such as PML-associated immune reconstitution inflammatory syndrome (PML-IRIS) with even worse prognosis.

For PML to occur, the replication of latent JCPyV must be reactivated and transmission to the brain must take place. As the vast majority of PML patients have some degree of underlying immunosuppression, the normal function of immune system seems crucial in regulating the latency and reactivation of JCPyV as well as transmission of neurotropic strains. Especially cell-mediated immunity, such as JCPyV-specific CD8+ CTLs, control the replication of neurotropic strains since low CTL levels are associated with worse PML prognosis (135). Replication of neurotropic JCPyV in the brain glial cells is a hallmark of PML onset. Although mutations in the viral genome seem to increase the ability of neurotropic strains to replicate and infect glial cells, viral mutations alone are not sufficient for the development of PML.

Symptoms and diagnosis

Clinical features of PML are determined by the anatomical location of demyelination. Demyelination of the brain white matter is progressive, reaching multiple centimeters in diameter, and it usually affects the frontal and parietal lobes of cerebral cortex, and the subcortical cerebrum (**Table 4**). In approximately half of PML patients, demyelination is also seen in the grey matter of the brain (136). Neuropathological features include oligodendrocytes with enlarged nuclei and giant, pleomorphic astrocytes with multiple nuclei; JCPyV is present in the nuclei of oligodendrocytes, and less frequently of astrocytes (137). PML can manifest as various neurological symptoms including motoric weakness, aphasia, cognitive dysfunction, headache, and gait abnormalities; less than 20% of patients experience sensory loss, visual impairment, diplopia, limb incoordination, or behavioral changes (138-140). PML is a life-threatening disease with poor prognosis; depending on the underlying disease approximately 30-70% die within months from diagnosis (138-140).

Although magnetic resonance imaging of the brain and clinical symptoms are cornerstones in PML diagnosis, either presence of JCPyV DNA in the cerebrospinal fluid by PCR or positive immunohistochemical staining of viral proteins in the brain tissue is needed for a definitive diagnosis (141). Given the fact that cerebrospinal fluid of approximately 40% of PML patients will remain negative for JCPyV DNA (138, 139), the JCPyV DNA negative PCR result from cerebrospinal fluid can alone neither confirm nor rule out PML. Due to the non-specificity of radiological and neurological findings, however, only the detection of JCPyV proteins in the brain confirms PML.

Table 4. Clinical, anatomical, and neuropathological features of PML.

Pathology
• Demyelination of frontal and parietal lobes of cerebral cortex, and the subcortical cerebrum
• Enlarged oligodendroglial nuclei and bizarre astrocytes with enlarged nuclei containing viral inclusion bodies
Clinical symptoms
• Motoric weakness, aphasia, cognitive problems, headache, gait abnormalities, sensory loss, visual impairment, seizures, diplopia, and limb incoordination
Diagnosis
• Clinical symptoms
• Magnetic resonance imaging of the brain showing demyelinated sites
• Positive JCPyV DNA PCR from cerebrospinal fluid or immunohistochemical confirmation of viral proteins in the brain biopsy

Abbreviations: JCPyV, JC polyomavirus.

Incidence

PML onset is associated with profound and long-lasting immunosuppression affecting especially cellular immunity. Few case reports have questioned whether underlying immunosuppression is a prerequisite for the development of PML (142). Today approximately 80% of PML cases are diagnosed in HIV-positive patients, 8.4% in patients with underlying lymphoproliferative disease, approximately 5% in transplant patients, and 0.44% in patients suffering from autoimmune disease (**Table 5**) (143). Even in these patients PML is extremely rare, a total of 31 PML cases with underlying hematological malignancies have been reported in Finland by 2016 (144). The highest PML incidence (1.3/1,000 person-years) is observed among HIV-1 patients (139) possibly related to the JCPyV promoter activating effect of HIV Tat protein (145) and simultaneous decrease of CD4+ and CD8+ T lymphocytes (135). Fortunately, the introduction of combined antiretroviral therapy (cART) has lowered the incidence of PML in HIV-positive patients but the prognosis is still quite poor with one-year survival rate of 33.33% (139).

Lymphoproliferative diseases most commonly associated with PML are chronic B-cell lymphocytic leukemia and Hodgkin lymphoma; PML-related mortality in these patients can reach up to 95.4% (146). PML is also observed after solid organ, especially kidney, transplantation (incidence: 1.24/1,000 person-years) with mortality due to PML approximately 41.2% (140). Markedly lower incidence (0.354/1,000 person-years) is observed among bone marrow transplant patients but the PML-related mortality is similar, 40% (140, 147). PML is rare among patients suffering from autoimmune diseases such as systemic lupus erythematosus and rheumatoid arthritis (incidence: 0.04/1,000 and 0.01/1,000, respectively) (143, 148); among these patients, however, the PML risk increases with prolonged use of certain biological drugs.

Table 5. Incidence of PML.

Clinical entity	Subgroup	Incidence (person-years)	1-year survival rate (%)
General population		0.044/1,000	Unknown
HIV		1.3/1,000	33.33
Transplantation	solid organ	1.24/1,000	58.8
	bone marrow	0.354/1,000	60
Lymphoproliferative disorders		Unknown	<10
Autoimmune disorders	RA	0.01/1,000	Unknown
	SLE	0.04/1,000	22

Abbreviations: RA, rheumatoid arthritis; SLE, systemic lupus erythematosus.

Biological drugs and PML risk

PML risk is elevated among patients treated with certain monoclonal antibodies including natalizumab, rituximab, efalizumab, and alemtuzumab (**Table 6**). These drugs are effective in treatment of autoimmune and lymphoproliferative diseases such as relapsing-remitting multiple sclerosis, Crohn's disease, severe plaque psoriasis, non-Hodgkin lymphoma, and rheumatoid arthritis. The emergence of PML in treated patients has however resulted in either restricted usage (natalizumab, rituximab, alemtuzumab) or temporary (natalizumab) or complete withdrawal (efalizumab) of these drugs.

Natalizumab, a humanized anti-α4-integrin antibody, prevents the invasion of lymphocytes to the inflammation site and increases the amount of circulating CD34-positive hematopoietic stem cells (149, 150). It has been approved as a monotherapy for severe relapsing-remitting multiple sclerosis and Crohn's disease (151, 152). JCPyV seropositivity, previous immunosuppressive treatment, and duration of natalizumab treatment increase PML risk: after two-year treatment of natalizumab, a ten-fold increase in risk (11.1/1,000) is observed among JCPyV seropositive patients with history of other immunosuppressive treatment (153). High JCPyV-specific antibody levels in natalizumab-treated patients without prior history of immunosuppression are associated with higher PML risk (154).

Rituximab, a monoclonal anti-CD20 IgG1 antibody, has been approved for the treatment of non-Hodgkin lymphoma and rheumatoid arthritis (155, 156). Rituximab-associated PML incidence is estimated to be approximately 1/20,000 and 1/10,000 for rheumatoid arthritis and non-Hodgkin lymphoma patients, respectively (147, 157). The proposed mechanism for rituximab-induced PML onset is similar to natalizumab: rituximab decreases the amount of circulating B lymphocytes and increases the release of CD34-positive cells from the bone marrow (158, 159).

Alemtuzumab has been approved for the treatment of B-cell chronic lymphocytic leukemia and multiple sclerosis (160, 161). It targets the surface protein CD52 abundant in B and T lymphocytes inducing their apoptosis (162). Four PML cases have been reported among alemtuzumab-treated B-cell chronic lymphocytic leukemia patients (163, 164) so monitoring for neurological symptoms is advised.

So far only efalizumab has been completely withdrawn from the market due to high PML incidence (1/400) in efalizumab-treated patients (165). No sufficient data has been accumulated on the PML risk among patients treated with newer biological drugs such as fingolimod and dimethyl fumarate.

Table 6. Monoclonal antibodies associated with increased PML risk.

	Natalizumab	Rituximab	Alemtuzumab
Indication	Severe RR-MS, Crohn's disease	non-Hodgkin lymphoma, RA	B-cell CLL, severe RR-MS
Target molecule	α4β1- and α4β7-integrin	CD20	CD52
Mechanism of action	Invasion of lymphocytes to inflammation site ↓, circulating CD34+ cells ↑	Circulating B lymphocytes ↓, circulating CD34+ cells ↑	Circulating lymphocytes ↓
PML incidence (treatment time)	1/1,000 (2 y) 11.1/1,000 (>2 y)	1/20,000 (>5 mo) 1/10,000 (unknown)	4 reported cases in B-cell CLL (3 mo)
Associated mortality (%)	22	90	100

Abbreviations: RR-MS, relapsing-remitting multiple sclerosis; RA, rheumatoid arthritis; CLL, chronic lymphocytic leukemia.

Immune reconstitution syndrome

Although case studies describing successful treatment of JCPyV infection in PML patients with chemokine receptor type 5 antagonist and $5HT_{2A}$ blocker have been described (166, 167), the reduction of immunosuppression is currently the only option to prevent the progression and improve the prognosis of PML. Normalization of cell-mediated immunity by withdrawal of immunosuppressive or immunomodulatory treatment may pose for the development of PML-IRIS. PML-IRIS is a T cell-mediated inflammatory state characterized by infiltration of JCPyV-specific CTLs to the brain parenchyma resulting in a massive destruction of JCPyV-infected glial cells and subsequent paradoxical worsening of neurological symptoms (168). PML-IRIS has an average mortality of 25% and 35% in natalizumab-treated multiple sclerosis and cART-treated HIV-positive patients, respectively (169, 170). Better outcome of PML-IRIS is achieved with treatment of corticosteroids that dampen the production of immune cells (169, 170).

1.4.2 Sequence variation of JC polyomavirus

NCCR

Mutations are frequently observed within the genome of JCPyV, especially in the NCCR. JCPyV can be classified into archetype strains with small point mutations within the NCCR and neurotropic strains having larger NCCR rearrangements. Archetype strains always have similar NCCR architecture that can be divided into arbitrary sequence blocks A (25 bp), B (23 bp), C (55 bp), D (66 bp), E (18 bp), and F (80 bp) (**Figure 7**) (65). Blocks A to F constitute the transcriptional control region (TCR) containing promoters and enhancer elements for early and late viral gene expression. Importantly, the NCCR determines tissue-specific gene expression of JCPyV due to the presence of multiple binding sites for host cell transcription factors such as Sp1 and NF-1 (171). Archetype strains are thought to cause asymptomatic primary infections since they are occasionally found in immunocompetent and immunosuppressed individuals without PML (65, 172-174). Sometimes minor deletions and point mutations occur

in the archetypal NCCR but the mutations seem not alter pathogenicity of such archetype-like strains (173, 174).

Within an individual, major rearrangements, duplications, and deletions may take place in the NCCR of an originally archetype virus leading to the development of neurotropic strains (175). Development of neurotropic strains within an individual is a prerequisite for PML onset (10) and these altered strains are detected in the brain and cerebrospinal fluid of PML patients (1, 63, 113, 176-178). Mutations within the NCCR are thought to alter the viral replication efficiency: neurotropic strains have been shown to replicate more efficiently in the brain glial cells than archetype strains (10). The enhanced replication of neurotropic strains could be due to the changes in the number of binding sites for the host cell transcription factors highly expressed in the glial cells, such as Tst-1, Pur-α, and YB-1 (69, 171). Especially the binding of transcription factors Spi-B and NF-1X to the NCCR seem to allow more efficient viral DNA replication and gene expression in the brain (179, 180).

Two theories on the underlying mechanism of the development of neurotropic strains currently exist. First, neurotropic strains may develop in the brain supported by the detection of neurotropic JCPyV in the brain of healthy individuals (177). Second, neurotropic strains could develop outside the brain for instance in the CD34-positive hematopoietic stem cells and B lymphocytes of the bone marrow and tonsils (121, 122). The V(D)J recombination machinery of B lymphocytes could allow accumulation of mutations within the NCCR (175). Finally, as neurotropic strains are occasionally secreted in the urine of non-PML patients (65, 172, 179, 180), the robust replication of archetype JCPyV in the kidneys may allow the emergence of neurotropic strains. More studies are needed, however, to establish the mechanisms behind the development of neurotropic strains.

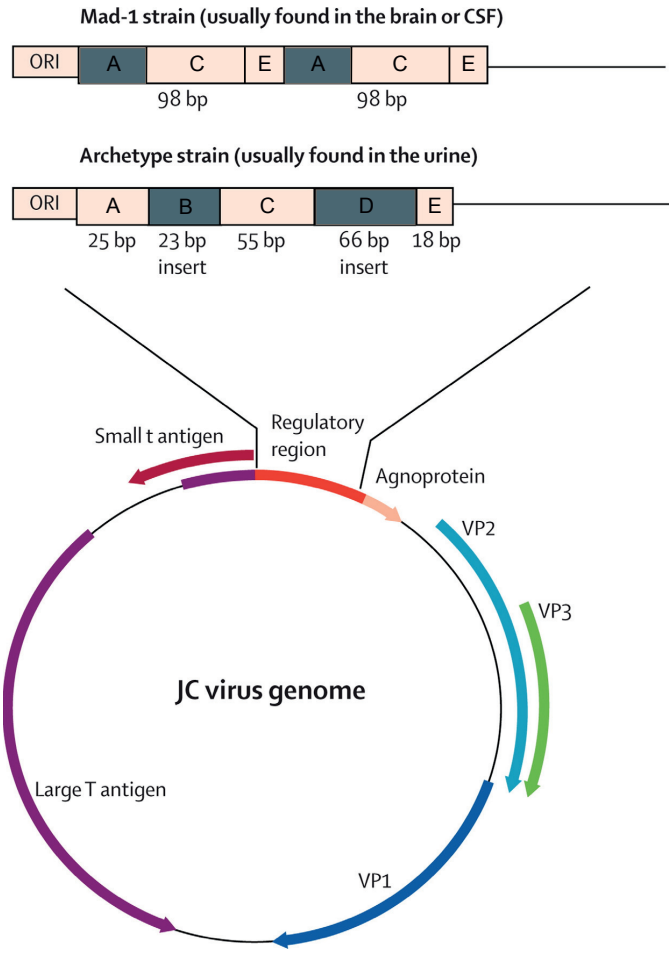


Figure 7. Structure of archetype and neurotropic JCPyV NCCR regions. The archetypal NCCR can be divided into blocks A (25 bp), B (23 bp), C (55 bp), D (66 bp), E (18 bp), and F. The NCCR of prototypic neurotropic JCPyV Mad-1 strain has duplications of A, C, and E blocks. Reproduced with permission from Tan and Koralnik 2010 (181).

Major capsid protein VP1

In addition to the mutations in the NCCR, neurotropic JCPyV strains often have mutations in the protein coding region, especially in the major capsid protein VP1 (**Table 7**); if these mutations are located in the receptor recognition region of VP1 they may affect the receptor specificity and subsequently the host cell tropism of JCPyV (7, 182). The most common mutations in neurotropic strains amplified in the brain tissue and cerebrospinal fluid of PML patients indeed affect amino acids in BC (55 and 60), DE (122), and HI loops (265, 267, 269, and 271) (7, 182). Especially mutations in the amino acids 269 and 55 are frequently detected (27% and 25%, respectively) in PML patients (182). These amino acids together with amino acid 267 may be critical for receptor binding as mutations in these amino acids inhibit hemagglutination and binding of JCPyV to the kidney cells and lymphocytes but not to the glial cells (7, 182).

Mutations in other viral genes

Mutations also occur in the highly conserved viral regulatory proteins that potentially affect the replication of JCPyV (**Table 7**). Neurotropic strains sometimes acquire point mutations affecting the helicase domain of LTag that might inhibit viral replication (183). Viral pathogenicity may also be affected by mutations in the agnoprotein and minor capsid protein genes as point mutations of VP2 and larger deletions of the agnoprotein have been observed in PML patients (184). The effects of such mutations on the viral pathogenesis still need to be confirmed in vitro.

Table 7. Amino acid mutations detected in neurotropic JCPyV strains and their proposed effect on viral pathogenicity. Only most frequently observed VP1 mutations are shown.

Viral protein	Mutation	Functional domain	Phenotypic effect in vitro
VP1	L55F K60N/E/M N265D/T S267Y/F/L S269F/Y/C	Host cell receptor recognition region	Alteration of host cell receptor binding
VP2	A6V G50R A69V A103V P174S R207G R322S A234T	Unknown	Unknown
Large T antigen	V392G	Helicase	Inhibition of viral replication
Agnoprotein	E59Q 21-aa-deletion (50…71)	Unknown	Unknown

Abbreviations: VP1, viral protein 1; VP2, viral protein 2; aa, amino acid.

1.5 Polyomavirus-associated nephropathy

1.5.1 BKPVAN results from uncontrolled replication of BK polyomavirus

BKPVAN results from reactivated replication of BKPyV persisting in the kidneys. Deterioration of immune system, especially of cell-mediated immunity, allows more robust replication of BKPyV in the proximal tubular epithelial cells of the kidneys leading to cellular detachment and necrosis characterized by interstitial inflammation, tubular atrophy, and fibrosis (185). BKPVAN is one of the major causes of allograft dysfunction and rejection after kidney transplantation and 1.1-10.3% of kidney transplant patients develop BKPVAN within a year from transplantation (186-189). Without any intervention, BKPVAN leads to rejection of kidney allograft in approximately 50% of kidney transplant patients within months from diagnosis (186, 188, 189). Only rarely BKPVAN has been diagnosed in association with other solid organ transplantations, and among kidney transplant patients the allogeneic graft seems to predispose to BKPVAN (190, 191).

Monitoring of BKPyV replication

No prophylactic or curative treatment for BKPVAN is available leaving the monitoring of kidney transplant patients as the only option for early intervention and prevention of BKPVAN. As prolonged high-level BKPyV viremia and viruria always precedes BKPVAN onset, the prospective screening for BKPyV replication in plasma and urine of kidney transplant patients every three months up to two-years post-transplantation and then annually until the fifth-year post-transplantation is recommended (**Figure 8**) (189). Robust BKPyV viruria (>7 log10 copies/mL) occurs in approximately 35-57% of asymptomatic kidney transplant patients and simultaneous urinary secretion of decoy cells, i.e. intranuclear viral inclusions, occurs in 13-30% of patients (188, 192-194). 7-13% of kidney transplant patients with viruria will proceed to have high-level BKPyV viremia (>4 log10 copies/mL) within 3 to 6 months post-transplantation (188, 192-194) which is a clear risk factor for BKPVAN (188, 195).

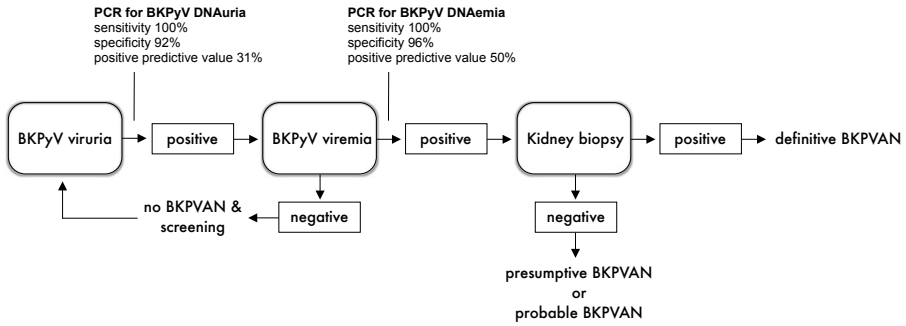


Figure 8. Screening of kidney transplant patients for BKPyV replication and BKPVAN. Presumptive and probable BKPVAN are characterized by high-level (>4 log10 copies/ml) and low-level (>3 log10 copies/ml) BKPyV viremia, respectively.

Diagnosis and intervention

BKPVAN can be presumed if high-level replication of BKPyV in both urine and blood is detected; for definitive diagnosis, however, a positive immunohistochemical staining of SV40 LTag or presence of typical cytopathic changes indicative of polyomavirus infection in the kidney biopsy is required (**Table 8**) (189). Prolonged viral replication in the kidneys leads to enlarged tubular cells with basophilic viral inclusions and sometimes parenchymal scarring due to progressive tubular atrophy and interstitial fibrosis of the kidneys (185). The diseased kidney tissue can be grouped into histological classes 1, 2, and 3 characterized by small cytopathic changes, broader inflammation, or larger tissue destruction, respectively (196). Due to extensive destruction of kidney tissue, classes 2 and 3 are associated with higher risk for allograft rejection (194). Due to patchy nature of viral inclusion bodies, sampling from multiple tissue sites is recommended to decrease the probability of false negative result (194).

Table 8. Diagnosis of BKPVAN. High-level BKPyV viruria is occasionally accompanied with urinary decoy cell shedding.

Laboratory finding	Details	Probable BKPVAN	Presumptive BKPVAN	Definitive BKPVAN
High-level BKPyV viruria	BKPyV DNA load of >7 log10 copies/mL	+	+	+
Low-level BKPyV viremia	BKPyV DNA load of >3 log10 copies/mL (in 2 measurements within 3 weeks)	+	+	+
High-level BKPyV viremia	BKPyV DNA load of >4 log10 copies/mL (in 1 of 2 measurements within 3 weeks)	-	+	+
Kidney biopsy	Viral cytopathic changes (class 1) Inflammatory infiltrates and tubulitis (class 2) Interstitial fibrosis and tubular atrophy (class 3)	-	-	+

Abbreviations: BKPyV, BK polyomavirus; BKPVAN, BK-polyomavirus associated nephropathy.

The first line treatment for BKPVAN is the reduction of immunosuppressive regimen, especially the reduction of antiproliferative drugs and calcineurin inhibitors (187, 189, 193). Approximately 15% of patients experience acute cellular rejection due to irreversible tissue damage and poor treatment response (187). Although antivirals such as cidofovir (197) and intravenous immunoglobulins (198) have been successfully used to treat BKPyV replication in kidney transplant patients, the results are controversial. Also, fluoroquinolones as prophylactic agents and vaccines have been developed (199). Finally, re-transplantation after allograft rejection is successful without evidence for BKPyV replication in majority (90%) of cases (200).

Risk factors

Several donor- and recipient-related factors have been suggested to increase risk for BKPVAN. First, a ten-fold risk is observed if a BKPyV seronegative recipient receives a graft from a seropositive donor (201). The risk for BKPyV viremia is also elevated if the recipient has low levels of anti-BKPyV antibodies (202). Risk factors include also male gender, older age (187), HLA (Human Leukocyte Antigen) mismatches, and previous acute rejection episodes (188).

The use of some immunosuppressive drugs has been associated with increased BKPVAN risk. For instance, the use of tacrolimus and corticosteroids as a maintenance therapy together with antithymocyte globulin as an induction therapy may increase the risk of BKPyV viremia and BKPVAN (188). It is unclear, however, whether use of a specific immunosuppressive drug or the overall state of immunosuppression plays more important role in the development of BKPVAN.

1.5.2 Sequence variation of BK polyomavirus

Similar to JCPyV, BKPyV can also be classified into archetype strains harboring a stable NCCR structure and mutated strains having a highly variable NCCR architecture.

Within the archetypal BKPyV NCCR, sequence blocks O (142 bp), P (68 bp), Q (39 bp), R (63 bp), and S (63 bp), respectively, can be distinguished (**Figure 9**) (203). The P, Q, R, and S blocks constitute the TCR containing promoters and enhancer elements for the expression of early and late viral genes (203). Archetype strains causing asymptomatic primary infections are mostly detected in immunocompetent individuals as well as in immunosuppressed individuals without BKPVAN indicating that they are not highly pathogenic (10, 172). The tissue-specific gene expression of BKPyV is determined by the NCCR containing multiple binding sites for host cell transcription factors expressed in the kidneys, such as Sp1, NF-1, NFAT, AP-1, and p53 (67, 204). Minor deletions and point mutations may occur in the NCCR but such mutations seem not alter viral pathogenicity. These strains are therefore called "archetype-like" and they seem not associate with the development of BKPVAN (10).

Within the organism, major rearrangements, duplications, and deletions in the NCCR may take place, particularly in the P, Q, and R blocks (205). The association of BKPyV NCCR mutations in the development of BKPVAN remains elusive; although mutated BKPyV NCCR regions are frequently detected in urine, plasma, and kidney tissue of BKPVAN patients (10, 122) and mutations seem to increase viral pathogenicity (10), it is unknown whether such mutations are auxiliary or necessary for the development of BKPVAN. Further, BKPyV strains with mutated NCCR are also found in asymptomatic kidney transplant patients without BKPVAN (206).

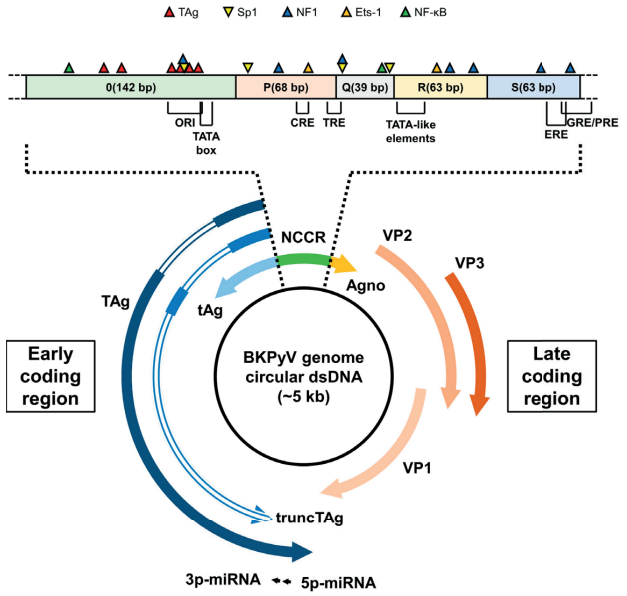


Figure 9. Structure of archetype BKPyV NCCR region. The NCCR can be divided into blocks O (142 bp), P (68 bp), Q (39 bp), R (63 bp), and S (63 bp). Further, the blocks P, Q, R, and S constitute the TCR containing promoters and enhancer elements for the viral gene expression. Reproduced with permission from Helle et al. 2017 (207).

The major capsid protein VP1

Active replication of BKPyV may sometimes lead to accumulation of mutations in the VP1: BKPyV strains with mutated VP1 are found in urine, plasma, and kidney tissue in approximately 40% of asymptomatic kidney transplant patients (118, 208-211). Enhanced viral replication in BKPVAN might allow accumulation of mutations

as most (40-75%) BKPVAN patients have BKPyV strains with VP1 mutations in their urine (118, 208-213). BKPVAN-associated VP1 mutations are often found in the BC loop, particularly in amino acids 69, 73, 75, 77, and 82, and less frequently in DE and EF loops (**Table 9**) (210-213). Mutations in the BC loop, especially in amino acids 62, 73, and 77, have been shown to alter the binding of BKPyV to the host cell receptor (208, 214). The mutation in the amino acid 77 of BC loop may also assist in evasion of host immune system (208). The importance of VP1 mutations in the development of BKPVAN remains to be established.

Mutations in other viral genes

Less is known about the impact of mutations in other viral genes in the development of BKPVAN. Occasionally mutations in the origin-binding domain of LTag are observed in BKPVAN-associated BKPyV strains (**Table 9**) (215). BKPVAN has not been associated with mutations located in the agnoprotein or minor capsid proteins VP2 and VP3 of BKPyV.

Table 9. Amino acid mutations detected in BKPyV strains associated with BKPVAN. Only most frequent VP1 amino acid mutations are presented.

Viral protein	Mutation	Functional domain	Phenotypic effect in vitro
VP1	K69R/H E73D D75N D77E E77Q E82D/N H139N D175E V201I	Host cell receptor recognition region	Alteration of host cell receptor binding
Large T antigen	Q150H	Origin binding domain	Unknown

Abbreviations: VP1, viral protein 1.

1.5.3 JC polyomavirus-associated nephropathy

Rarely, nephropathy may be caused by lytic infection of JCPyV in the kidneys; to date, a total of fourteen cases have been described (3, 4). Approximately 20% of kidney transplant patients experience high-level JCPyV viruria (>7 log10 copies/mL) and simultaneous low-level (<3 log10 copies/mL) JCPyV viremia is rarely detected (3, 4, 216). In association with JCPVAN, the viremia is also very low or absent (3, 4, 217). Currently monitoring of kidney transplant patients for JCPVAN is not recommended due to rarity of the disease (218). High urine JCPyV loads accompanied with histopathological changes characteristic for nephropathy and a positive JCPyV LTag staining of the kidney tissue are cornerstones in JCPVAN diagnosis. Reduction of immunosuppression is used for treatment (217, 219). The exact predisposing factors for JCPVAN are unknown, but similar to BKPVAN, male gender, acute rejection episodes, and recipient seronegativity seem to increase the risk (4).

1.6 Massive parallel sequencing

1.6.1 Novel high-throughput sequencing techniques

Massive parallel sequencing (MPS) includes novel next-generation sequencing (NGS) techniques available from 2005 onwards. They have for the most part replaced the traditional Sanger sequencing, a first-generation sequencing technique that uses a chain termination method for sequencing, and electrophoresis for amplicon separation (220). Advantages of Sanger sequencing include a very low error rate (0.3%) and a relatively long sequence read length (~900 bp). It is, however, slow and expensive: the sequencing of first human genome by Sanger sequencing took over a decade and cost approximately 3 billion dollars. Today, the same can be achieved with MPS techniques faster and with only fraction of the cost (221). MPS techniques differ in their need for pre-amplification, sequencing chemistry, and base pair detection technique (**Table 10**) (11). The basic principle is nevertheless the same: the sequencing of clonally amplified or single DNA molecules is done simultaneously in thousands to billions of parallel, spatially separated sequencing reactions producing thousands to millions of short sequence reads in a single machine run.

All MPS techniques consist of library preparation, sequencing, and data analysis steps. Based on the length of sequencing read, MPS techniques can be classified into short-read and long-read sequencing techniques with maximum sequence read lengths of up to 600 bp and 900,000 bp, respectively (**Table 10**). MPS techniques have revolutionized the field of human genetics and are routinely used to detect genetic variation associated with various clinical manifestations including single nucleotide point mutations, indels (insertion-deletions), structural variations, and copy number alterations. Whole genome sequencing allows the investigation of all mutations within an individual genome while whole exome sequencing is used to study protein coding region mutations associated with rare inherited disorders (reviewed in (222, 223)). MPS techniques have also proven useful in transcriptomics and epigenomics of all eukaryotes and they are employed to study DNA-protein interactions by chromatin immunoprecipitation sequencing (reviewed in (224-226)). MPS techniques allow detailed characterization of genomic organization and biological functions taking place within a single cell. Further, single cell sequencing of circulating tumor cells has enabled investigation of their role in the metastasis formation of for instance lung cancer (227).

Table 10. Comparison of first-generation Sanger sequencing and next-generation sequencing technologies. Of note, the duration of sequencing, read length, and maximum output depends on the version used.

	Platform	Sequence by	Run time	Maximum read length (bp)	Maximum output/run (Gb)	Error rate (%)	Sources of error
Short-read	Sanger	Chain-termination	20 min-3 h	900	0.0021	0.3	N/A
	Illumina	Synthesis	1-3 d	125-300 (PE)	7.5-6,000	<1	Substitutions, AT and GC biases
	SOLiD	Synthesis	10 d	50 (PE)	240	0.01	AT bias
	Ion Torrent	Synthesis	2-7 h	200-600 (SE)	2-25	1	Indels, homopolymers
Long-read	Pacific Biosciences	SMRT	30 min-20 h	15,000-30,000	20	~1	Indels
	Oxford Nanopore Technologies	SMRT	1 min-64 h	Up to 900,000	315	15	Indels, homopolymers

Abbreviations: bp, base pair; Gb, gigabase; N/A, not available; d, days; PE, paired-end; SE, single-end; SMRT, single-molecule real-time sequencing.

Short-read sequencing

Short-read sequencing techniques produce sequence reads with a size of few hundred base pairs. All short-read sequencing techniques use pre-amplification step to increase the signal-noise ratio as they are not sensitive enough to detect signal from an individual template molecule. Short-read sequencing techniques differ in their pre-amplification method (solid-phase or emulsion PCR), sequencing chemistry (cyclic reversible termination or ligation) and base pair detection technique (fluorescence or voltage change) (**Table 10**, **Figure 10**). First, Illumina sequencing uses cyclic reversible termination with fluorescent labeled nucleotides. Repeat regions are problematic due to short sequence read length of Illumina sequencing, and reverse terminators and modified polymerases produce GC bias (i.e. difference in the GC content between sequence reads and reference read) and substitution errors (228). Second, the detection of Ion Torrent sequencing is based on the release of a proton after incorporation of a base and it is prone to substitution (0.1%) and homopolymer (3.5%) errors (229). Sequencing by ligation is used only by SOLiD sequencing where DNA polymerase has been replaced by DNA ligase; although being prone to AT bias, SOLiD sequencing can distinguish sequence polymorphism from sequencing error as each position is probed twice (230).

Advantages of short-read sequencing techniques include low error rate, fast processing, and high-throughput. The high accuracy of short-read sequencing is due to the fact that low base quality and accuracy can be compensated by increasing the coverage of a single base (231). Problems include the requirement for the error-prone pre-amplification step, the need for large amount of original material (3-20 µg), GC and AT biases as well as problems in sequencing of repeats and homopolymers (11).

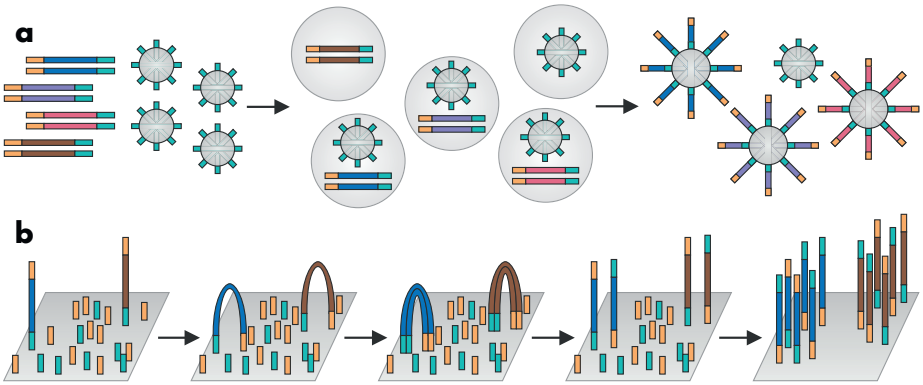


Figure 10. Pre-amplification strategies of different short-read sequencing techniques. Emulsion PCR is used by Ion Torrent sequencing where each DNA molecule is amplified within a vesicle containing all reagents for amplification (a). Illumina sequencing uses bridge-amplification for amplification of DNA molecules which results in clusters of amplicons originating from a single template molecule (b). Reproduced with permission from Shendure et al. 2008 (232).

Single-molecule real-time sequencing

Long-read sequencing techniques by Pacific Biosciences (PacBio) and Oxford Nanopore have many advantages over the short-read sequencing techniques: both have capacity to sequence very long reads (>10,000 bp), they require far less starting material (<1 µg), and do not use the error-prone pre-amplification step. Long-read sequencing is used for instance in de novo sequencing of microbes (233), and to study structural variation of human genome such as copy number variation (234, 235).

Long-read sequencing techniques use single-molecule real-time (SMRT) sequencing, where DNA polymerization can be visualized real-time in a base-pair resolution (236, 237). The PacBio SMRT sequencing is performed in thousands of SMRT wells with zero-mode waveguide detectors for the detection of incorporated fluorescent nucleotides (**Figure 11**) (238). Before sequencing, hairpin adapters are ligated to both ends of a template molecule to generate a SMRTbell™ template; this allows the polymerase to sequence single template molecule multiple times without interruptions thereby increasing the sequence coverage (239). Pacbio SMRT sequencing has very low signal-noise ratio, approximately 25:1 (236), and although the accuracy of PacBio SMRT sequencing is quite low (83%), it can be increased to approximately 99% by sequencing the template multiple times (236). As errors are randomly distributed and do not accumulate at the ends of sequence reads, the consensus accuracy can be as high as 99.99% (240). In another SMRT technique, Oxford Nanopore sequencing, the template molecule is allowed to move through nanometer-sized holes producing a nucleotide-specific voltage (237, 241). Even templates up to hundreds of thousands of base pairs in size can be sequenced by Oxford Nanopore sequencing as it does not rely on a polymerase (242). The error rate is very high, however, approximately 15% for Minion sequencer (242, 243): especially indels, substitutions, and homopolymers pose a problem for Oxford Nanopore sequencing (235).

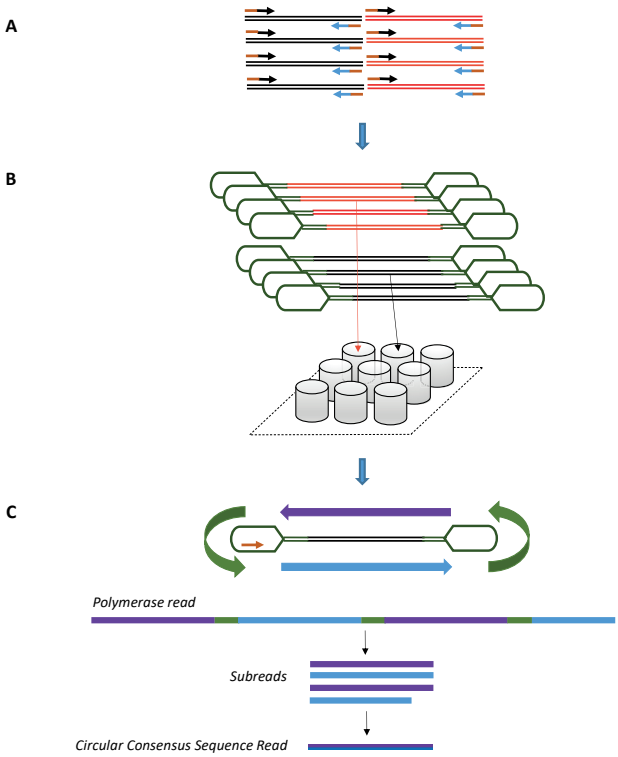


Figure 11. Schematic workflow of PacBio SMRT sequencing. First, amplicons are provided with PacBio barcodes (a) followed by library preparation where hairpin adapters are ligated to the both ends of an amplicon molecule to generate a circular SMRTbellTM template (b). During sequencing, the polymerase reads a single template molecule multiple times generating a polymerase read consisting of multiple subreads that can be used to produce a circular consensus sequence read (c).

1.6.2 Library preparation, data analysis, and pitfalls in sequencing

Both short- and long-read sequencing techniques require a template processing step prior sequencing. Library preparation consists of mechanic or enzymatic fragmentation of template molecule, cleaning, end repair, and adapter ligation steps. Adaptors may include molecular barcodes to allow sequencing of different samples in a single flow cell, universal PCR primers, hybridization sequences for the binding of template to the surface of the well, and recognition sites for the polymerase. To ensure data quality and correct interpretation of the results, special care is required in all these steps. Challenge during library generation is to increase the amount of specific target sequence without simultaneously increasing the number of unspecific products, such as host-derived DNA.

After sequencing, the raw sequence data has to be modified in order to obtain biologically meaningful results. Data analysis includes many bioinformatics processes such as filtering of uninformative reads in order to reduce the number of low-quality reads, and removal of adapter sequences and duplicate reads to improve the accuracy of sequence alignment (244). Typically, data analysis also includes steps such as quality analysis, base calling (i.e. conversion of image data into sequence reads),

sequence read mapping, variant identification, and annotation (245). Many factors such as primers, nucleic acid extraction, library preparation, and PCR can affect the quality of sequence reads (246) which needs to be taken into account during the data analysis. Errors in sequence data can be generated by signals from nearby clusters during sequencing and dephasing (247). High coverage decreases the probability of misinterpretation of the results due to sequencing errors and for instance 60-20-fold coverage is routinely used in Illumina sequencing (247).

1.6.3 Massive parallel sequencing in virological diagnostics

MPS has increased the diagnosis of previously undiagnosed infectious diseases and is currently being adapted into routine virological diagnostics. High coverage allows the investigation of viral populations and low-abundance variants in microbial samples (38). Targeted sequencing of smaller genomic regions using short-read sequencing seems feasible in routine diagnostics as it is cheaper and less time-consuming than long-read sequencing. However, contaminating host sequences decrease the sensitivity of detecting a specific pathogen. To overcome this, various targeted enrichment methods, such as probe-based enrichment, are used to increase the amount of target sequence. MPS can be used to detect minor populations of drug-resistant HIV strains and to identify pathogens in diseases of unknown origin (12, 248).

Currently used diagnostic tests such as PCR rely on the prior knowledge of a pathogen. Although MPS-based diagnostic tests currently available cannot yet provide as fast turn-around time as real-time PCR, they have decreased dramatically the amount of original material needed for pathogen detection and enabled studies of non-cultivable microorganisms. One advantage of MPS-based diagnostic tests is the ability to simultaneously analyze several genomic regions in a single test. Faster library preparation and sequencing, more user-friendly data analysis, and clear guidelines for validation are, however, needed in order to improve the feasibility of MPS in virological diagnostics. MPS-based diagnostic tests are currently being implemented in many diagnostic laboratories worldwide.

1.6.4 Ethical considerations

Many ethical issues need to be considered before MPS techniques can be implemented in routine diagnostics. The huge amount of sequence data generated by MPS techniques requires computational power for storage, management, and data processing. A lot of the obtained information is unintended, and that raises concern about how and whether to use the data or store them for future purposes. European clinical laboratories have agreed that only data that has been cleaned from human sequences can be shared. Importantly, data sharing is crucial in order to improve bioinformatics tools used in the data analysis and to ensure similar results between different laboratories. Also an issue of discussion is the balance between high quality data and reasonable sequencing costs which are strongly affected by higher sequencing coverage and depth. The cost of MPS will continue to drop in the future and the cost reduction should be weighed against the quality of a sequence as higher coverage does not necessarily mean better quality of a sequence data. Despite ethical challenges associated with the use of MPS in diagnostics, the high-throughput and accuracy of MPS techniques might enable earlier identification of a pathogen and increase the probability of a correct diagnosis, issues that will ultimately improve patient monitoring.

2 AIMS

The aim of this thesis was to uncover how JCPyV and BKPyV sequence variation contributes to disease pathogenesis. The detailed characterization of viral populations associated with the development of polyomavirus-associated diseases, such as PML and polyomavirus-associated nephropathy, has been partly hindered by the lack of sufficiently sensitive detection methods. Here using high-throughput MPS techniques, the role of mutations within the viral genome in the pathogenicity of JCPyV and BKPyV and in the development of PML, BKPVAN, and JCPVAN was studied.

The more specific aims were:

- To characterize PML-associated mutations originating from individual neurotropic JCPyV strains, with emphasis on the NCCR, VP1, and LTag (I),

- To study whether mutations within the viral genome, especially in the NCCR, VP1, and LTag, are necessary for the development of JCPVAN (II),

- To investigate the role of BKPyV TCR mutations in the development of BKPVAN (III, IV), and

- To evaluate the potential of BKPyV TCR mutations as biomarkers for BKPVAN (III, IV).

3 EXPERIMENTAL DESIGN

This section summarizes the sample material, methods, and clinical characteristics of patients used in this thesis. The reader is kindly instructed to have a look at the respective original publications (I-IV) for more detailed description of experimental design.

3.1 Sample material and clinical characteristics of patients

3.1.1 PML patients (I)

Cerebrospinal fluid samples from three adult PML patients (median age: 74, range: 59-82, one female, two males) were analyzed in this work (**Table 11**). All patients fulfilled the criteria of definitive PML (141). Two patients had an underlying immunosuppressive disease; diffuse large B-cell non-Hodgkin lymphoma (patient 1) and follicular stage IVA lymphoma (patient 2). Both patients 1 and 2 had received rituximab treatment prior to PML diagnosis (14 and 30 courses, respectively). Patient 3 had no prior immunosuppressive medication or condition apart from type 2 diabetes.

All patients demonstrated neurological symptoms characteristic to PML. Severe headache accompanied with vision disturbances were observed in patient 1 and brain magnetic resonance imaging revealed lesions in the right and left occipital lobes, and right frontal subcortical lobes. High amount of JCPyV DNA was detected in the cerebrospinal fluid by quantitative PCR (qPCR). Patient 1 died 14 months after PML diagnosis. Patient 2 had neurological symptoms including tiredness, apraxia of both hands, and visual impairment. White matter lesions in the temporo-occipital region and high-level JCPyV replication in the cerebrospinal fluid was detected and the patient died one month post-diagnosis. Patient 3 developed first dysarthria and delirium, and brain computed tomography showed infarct in the right frontal lobe. The symptoms rapidly progressed into epileptic seizures, cognitive decline, and speech apraxia accompanied with lesions in the both frontal lobes, right occipital lobe as well as both sides of the cerebellum and pons. High amount of JCPyV DNA was detected in the cerebrospinal fluid. Patient 3 died one month after the diagnosis.

Table 11. Clinical characteristics of PML patients.

Patient	Gender, age (years)	Underlying disease	Immuno-suppression (no. of cycles)	Symptoms	Anatomic location of lesions	Outcome
1	Female, 74	Diffuse large B-cell non-Hodgkin lymphoma	Rituximab (14)	Headache, vision disturbances	Occipital and frontal lobes	Deceased (14 mos.)
2	Male, 82	Follicular stage IVA lymphoma	Rituximab (30)	Tiredness, apraxia, visual impairment	Temporal and occipital lobes	Deceased (1 mo.)
3	Male, 59	None	None	Dysartria, delirium, epileptic seizures, cognitive decline, speech apraxia	Frontal and occipital lobes, cerebellum, pons	Deceased (1 mo.)

3.1.2 Kidney transplant patients with JCPyV viruria (II)

To study the role of mutations within JCPyV genome in severe kidney diseases, urine samples of twenty patients who had undergone a kidney transplantation at Helsinki University Hospital were collected (**Table 12**). Except for two pediatric patients aged 17 and 2, all patients were adults with mean age of 49 years (range: 18–73). A great majority (~88,9%) of the patients were males. The immunosuppressive therapy consisted of cyclosporine or tacrolimus, mycophenolate and steroids, and patients with higher immunological risk received also an induction therapy with basiliximab. All patients had stable allograft function with the mean plasma creatinine level of 133 µmol/L (range: 53–268) at sampling. All patients were alive with functioning allograft at the end of mean follow-up period of 26 months and average plasma creatinine level of 139 µmol/L (range: 53–268).

Table 12. Clinical characteristics of stable kidney transplant patients with JCPyV viruria.

Patient	Gender	Age at KTx (years)	Time from KTx at sampling (days)	Plasma creatinine at sampling (µmol/L)	Symptoms at sampling	Plasma creatinine at the end of follow-up (µmol/L)
4	Female	65	+11	149	None	104
5	Female	67	+179	141	None	234
6	Male	46	+363	131	None	131
7	Male	18	+9	104	None	160
8	Male	71	+376	161	None	169
9	Male	48	+204	119	None	116
10	Male	47	+13	235	None	235
11	Male	19	+5	123	None	123
12	Male	52	+194	159	None	147
13	Female	56	+11	149	None	85
14	Male	21	+223	185	Minor graft dysfunction	152
15	Female	32	+11	70	None	112
16	Male	73	+11	202	None	149
17	Male	62	+2707	106	None	107
18	Male	62	+574	167	None	268
19	Female	17	+2176	65	None	70
20	Male	2	+719	42	Pollakisuria, pyuria	53
21	Male	47	+231	118	None	131
22	Male	49	+11	134	None	131
23	Male	50	+299	96	None	100

Abbreviations: KTx, kidney transplantation.

Also included were four urine and three plasma samples from one kidney transplant patient with an extremely rare histologically confirmed JCPVAN (**Table 13**). Samples were taken during the follow up period of 28 months: all, except plasma samples 1 and 2, were taken after the diagnosis of JCPVAN. The patient has been described in detail elsewhere (219). Briefly, at age of 15 years the patient underwent a transplantation at Helsinki University Hospital receiving a cadaveric kidney allograft.

At 11 months post-transplantation, the allograft function deteriorated (plasma creatinine level ~500 µmol/L) and the kidney biopsies revealed intensive interstitial inflammation and tubulitis. The patient had high-level (>1.0E+08 copies/mL) JCPyV viruria but extremely low JCPyV viremia (<400 copies/mL). BKPyV replication was not detected in plasma or urine. The diagnosis of JCPVAN was based on the positive SV40 LTag immunoperoxidase staining of multiple biopsies and positive JCPyV DNA but negative BKPyV DNA by PCR. Retrospectively the patient was shown to be JCPyV IgG seronegative before transplantation but developed both IgM and IgG responses after transplantation; anti-BKPyV antibodies remained negative at all times. Despite reduction of immunosuppression, the kidney allograft was lost 27 months after the transplantation due to the primary JCPyV infection and chronic antibody-mediated rejection.

Table 13. Samples of the kidney transplant patient with histologically confirmed JCPVAN.

Sample	Source	Time from KTx (days)	Time from JCPVAN diagnosis (days)	Histological finding (Banff criteria)	Anti-SV40 LTag staining
1	Plasma	+271	-54	N/A	N/A
2	Plasma	+278	-46	N/A	N/A
3	Plasma	+338	+14	Interstitial nephritis, borderline acute rejection (1)	Positive
4	Urine	+341	+17	Interstitial nephritis, borderline acute rejection (1)	Positive
5	Urine	+555	+231	Interstitial fibrosis and tubular atrophy, no rejection	Negative
6	Urine	+688	+364	Interstitial fibrosis and tubular atrophy, no rejection	Negative
7	Urine	+842	+518	Glomerular sclerosis, interstitial fibrosis and tubular atrophy, antibody-mediated rejection	Negative

Abbreviations: KTx, kidney transplantation; JCPVAN, JC polyomavirus-associated nephropathy; N/A, not available.

3.1.3 Kidney transplant patients with BKPyV viremia (III, IV)

To assess the role of BKPyV TCR variation in the development of BKPVAN, plasma samples from nine kidney transplant patients from Helsinki University Hospital (2-12 samples per patient, 53 in total, **study III**) and pairs of urine and plasma samples from fifteen kidney transplant patients from Medical University of Vienna (**study IV**) were included (**Table 14**). Most (70.8%) patients were males and the median age was 58 years (range: 23–72). After transplantation, the patients received either a maintenance therapy with mycophenolate, methylprednisolone, and tacrolimus or cyclosporine (**III**), or they were treated with mycophenolate, corticosteroids, tacrolimus or cyclosporine A (**IV**). An induction therapy with anti-thymocyte globulin or basiliximab was administered for patients 11, 12, 14-21, 23, and 24 (**IV**). BKPyV viremia (>1.0E+03

copies/mL) was detected with median of 107 days post-transplantation (range: 0–468) and was successfully treated in all patients with reduction of immunosuppressive regimen. Nine patients experienced allograft dysfunction indicated by a decrease in the estimated glomerular filtration rate. All patients were alive with functioning allograft at the end of mean follow-up of 1,082 days (range: 131–2,776); patients 6 and 24 lost the allograft due to antibody-mediated rejection. Two kidney transplant patients with stable allograft function and no BKPyV viremia or viruria detected in routine screening were included as controls in study III.

Table 14. Clinical characteristics of kidney transplant patients with BKPyV viremia. Only the sampling date for the sample used for characterization of BKPyV TCR is shown.

Study	Patient	Gender	Age at KTx (years)	Time from KTx at sampling (days)	Kidney status at sampling	Duration of BKPyV viremia at sampling (days)	Kidney status at the end of follow-up
III	1	Male	67	+141	Stable	55	Stable
	2	Male	61	+90	Stable	0	Stable
	3	Male	72	+102	Stable	0	Stable
	4	Female	47	+433	Stable	219	Stable
	5	Male	54	+190	Graft dysfunction	0	Stable
	6	Male	45	+148	Stable	47	Rejection
	7	Male	69	+166	Graft dysfunction	9	Stable
	8	Male	45	+202	Stable	0	Stable
	9	Female	71	+140	Stable	0	Stable
IV	10	Male	58	+140	Stable	0	Stable
	11	Male	59	+159	Stable	77	Stable
	12	Male	23	+468	Stable	0	Stable
	13	Male	72	+207	Stable	71	Stable
	14	Male	52	+40	Stable	0	Stable
	15	Male	27	+126	Stable	28	Stable
	16	Male	60	+303	Graft dysfunction	0	Stable
	17	Male	59	+100	Graft dysfunction	28	Stable
	18	Female	45	+127	Stable	22	Stable
	19	Female	39	+108	Stable	78	Stable
	20	Female	67	+90	Graft dysfunction	90	Stable
	21	Male	69	+225	Graft dysfunction	14	Stable
	22	Female	58	+224	Graft dysfunction	131	Stable
	23	Male	44	+235	Graft dysfunction	148	Stable
	24	Female	49	+69	Graft dysfunction	14	Rejection

Abbreviations: KTx, kidney transplantation; BKPyV, BK polyomavirus.

3.1.4 Ethical permission

The use of human plasma, urine, and cerebrospinal fluid samples in these studies was approved by the Coordinating Ethical Committee of the Helsinki and Uusimaa Hospital District (54/13/03/00/2015) and by the ethics committee of the Medical University of Vienna (EK2064/2016).

3.2 Total nucleic acid extraction (I-IV)

Total nucleic acids were extracted using either the automated NucliSENS EasyMAG extraction platform (bioMérieux SA, Marcy l'Etoile, France) (**IV**) or the automated Magna Pure LC extractor (Roche Molecular Systems, Pleasanton, CA) (**II**, **III**) according to the manufacturer's instructions. For the amplification of the complete JCPyV genome from the cerebrospinal fluid of PML patients (**I**), nucleic acids were extracted using the Magna Pure LC 1.0 extraction platform, and nucleic acids from the urine of kidney transplant patients with BKPyV viremia (**IV**) were extracted using the Qiagen viral RNA mini kit (ThermoFisher Scientific, Waltham, MA) according to the manufacturer's instructions. For the extraction, either 200 µl (cerebrospinal fluid), 100-200 µl (urine), or 200 µl (plasma) of original sample material was used. From six patients either 100 µl (14, 16, 18) or 80 µl (12, 13, 15) of original plasma was available (**IV**). The extracted nucleic acids were eluted in 25-50 µl (cerebrospinal fluid), 60-100 µl (urine), or 25-50 µl (plasma) of elution buffer and stored in -20°C until analyzed.

3.3 Quantitative Taqman real-time PCR (I-IV)

Urine and plasma samples of kidney transplant patients (**II-IV**) and cerebrospinal fluid samples of PML patients (**I**) were screened for the presence of JCPyV and BKPyV replication using the previously described Taqman qPCR (249) with some modifications. The qPCR reaction consisted of either 1x Taqman universal master mix (Applied Biosystems, Life Technologies, Waltham, MA), 900 nM of both primers, 175 nM of fluorescent-labelled TAMRA probe (Integrated DNA Technologies, Iowa, USA), and 10 µl of template in a total reaction volume of 50 µl (**I-III**) or 5 µl of template was added to the amplification mixture using TaqMan universal PCR master mix (PE Applied Biosystems, Waltham, MA) in a total reaction volume of 25 µl (**IV**). The reactions were incubated either in 50°C for 2 min, 95°C for 10 min followed by 45 cycles of 15 sec at 95°C and 1 min at 60°C (**I-III**) or 50°C for 3 min, 95°C for 10 min followed by 45 cycles of 95°C for 15 sec, 55°C and 72°C each for 30 sec (**IV**).

3.4 Characterization of complete JCPyV genomes (I, II)

3.4.1 Amplification of complete viral genome

Complete JCPyV genomes were amplified from cerebrospinal fluid of PML patients (**I**), and urine and plasma of kidney transplant patients with JCPyV viremia (**II**). In order to amplify every nucleotide of the genome, the primers were designed to the adjacent regions within the LTag gene. The PCR reactions contained 1x Q5 reaction buffer (New England Biolabs, Ipswich, MA), 200 µM dNTP mix, 500 nM of both primers, 0.5 U of Q5 Hot Start High-Fidelity DNA polymerase (New England Biolabs, Ipswich, MA), 1 µl of template, and sterile water in a 25-µl total reaction volume. Following conditions were used for amplification: 98°C for 30 sec followed by 35 cycles of 98°C for 10 sec, 64°C for 30 sec, and 72°C for 3 min with final extension of 2 min at 72 °C. Amplification using primers with PacBio SMRT sequencing barcodes was performed using the same reaction conditions. The product quality was confirmed on 0.8% agarose gel.

3.4.2 Single-molecule real-time sequencing and sequence analysis

The SMRT library preparation and sequencing were done according to the manufacturer's instructions (Pacific Biosciences, Menlo Park, CA). Briefly, contaminants were removed and DNA damage including nicks, apurinic/apyrimidinic sites, and oxidation damage were repaired. For library generation, blunt ends were generated at both ends of each amplicon followed by the ligation of hairpin-loop adapters to create a SMRTbell™ template. Failed ligation products were removed by exonuclease treatment. Sequencing was performed using PacBio RS II System and C4 chemistry. For sequence analysis, tools available in the SMRT Analysis Software (version 2.3.0) were utilized: circular consensus sequences were generated using the Long Amplicon Analysis tool with default parameters. The results were confirmed with polymerase reads containing 5 or more full passes of complete JCPyV genome using the Reads of Insert Mapping tool with modified parameters (minimum predicted accuracy: 80%, minimum read length of insert: 4,900 bp).

3.4.3 Characterization of complete JCPyV sequences

The classification of patient strains was based on the architecture of NCCR using JCPyV CY strain as a reference archetype strain (study I: AB038249; study II: M35834). Genotyping was based on the comparison of coding region sequences of patient strains with the respective prototype genotype strains (110). For more detailed analysis, the coding region sequences of genotype 1A (AF015526), 1B (AF015527), 2B (AF015533), and 4 (AF015528) were used. Amino acid sequences were generated using the EMBOSS Transeq tool (European Bioinformatics Institute, European Molecular Biology Laboratory) with default parameters. In the Results, the nucleotide numbering of NCCR and coding region is according to the archetype JCPyV CY strain and the corresponding genotype strain, respectively.

3.5 Characterization of transcriptional control region of BKPyV (III, IV)

3.5.1 Amplification of BKPyV TCR

The BKPyV TCR region was amplified from plasma of nine kidney transplant patients (**III**) and from pairs of urine and plasma samples of fifteen kidney transplant patients (**IV**). Each PCR reaction consisted of 1x PCR buffer I (Thermo Fisher Scientific, Waltham, MA), 0.2 mM dNTP mix, 0.5 µM of both primers, 0.625 U of Amplitaq Gold DNA polymerase (Thermo Fisher Scientific, Waltham, MA), 1-3 µl (**III**) or 10 µl (**IV**) template, and sterile water in a final reaction volume of either 25 µl (**III**) or 50 µl (**IV**). The aim was to sequence complete BKPyV TCR (~300 bp) in a single continuous read. The reactions were incubated at 95°C for 10 min followed by 35 cycles of 95°C for 15 sec, 67°C for 30 sec, 72°C for 25 sec, and finally 5 min at 72°C. The quality of products was confirmed on 1.75% agarose gel.

3.5.2 Targeted short-read sequencing

The sequencing libraries were constructed according to the manufacturer's instructions (Illumina Inc., San Diego, CA). The fragments were polished, A-tailed, and ligated to a TruSeq truncated adapter. Purification of ligation reactions was done using AMPure XP beads (Agencourt, Beckman Coulter). PCR of the libraries was done using Phusion Hot Start II DNA Polymerase (Thermo Fisher Scientific, Waltham,

MA), Index P7 primers, and full-length P5 adapter primers. The reactions were pooled and purified with AMPure XP beads. Size selection of the pool was done as previously described (250). The obtained library pool was paired-end sequenced on a MiSeq Sequencer using the v3 600 cycle kit. Sequencing was done in a paired-end manner (R1:326 bp; R2:286 bp) to obtain sequence information from both forward and reverse strands of an amplicon.

3.5.3 Sequence analysis

The in-house pipeline used for the sequence analysis is depicted in **Figure 12**. Shortly, adapter sequences were removed from read pairs and reads were trimmed using the cutadapt tool with parameters for minimum sequence quality (Phred score 30) and minimum read length (50 bp). Read pairs containing low-quality N-bases were discarded. The FLASH tool was used to combine read pairs containing both forward and reverse primer sequences. For clustering, the DNACLUST tool with identity cut-off of 97.5% was used. The results were confirmed by mapping using the BWA-MEM mapping tool and the Tablet tool was used for alignment visualization. The cut-off value for cluster size was set at 1,000 sequences comprising >90% of all sequences per sample. The cluster reference sequences were used to characterize BKPyV TCR regions present in a sample using the TCR of archetype BKPyV WW strain as a reference (203). Nucleotide numbering in the Results is according to the TCR sequence of the archetype BKPyV WW strain (AB211371.1) (nucleotides 35…267).

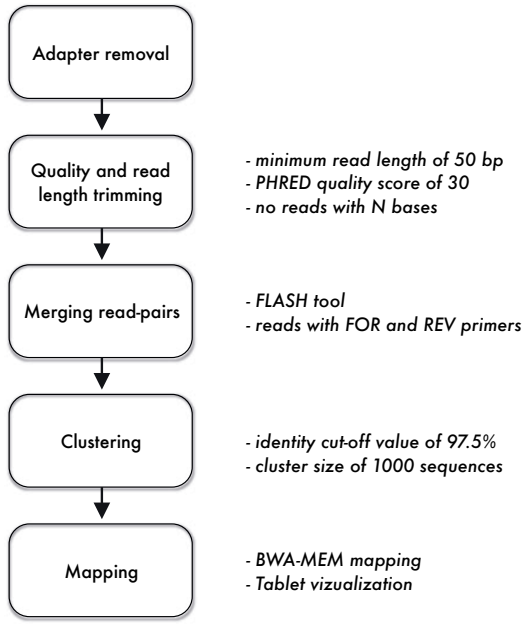


Figure 12. Overview of the bioinformatics pipeline used for the characterization of BKPyV TCR.

3.6 BKPyV microRNA analysis (III)

Total RNA was extracted from 620 µl of original plasma and eluted in 95 µl using the mirVana PARIS RNA and Native Protein Purification Kit (Ambion, Thermo Fisher Scientific, Waltham, MA) according to the manufacturer's instructions. To control reverse transcription and microRNA amplification, the extractions were spiked with 25 fmol of Caenorhabditis elegans cel-39-3p miRNA (Integrated DNA Technologies, Coralville, IA). The reverse transcription and microRNA amplification were performed using Taqman miRNA assays (Thermo Fisher Scientific, Waltham, MA) targeting bkv-miR-B1-5p, and bkv-miR-B1-3p, and cel-miR-39-3p. The reverse transcription reaction contained 10 ng of total RNA, 1x reverse transcription buffer, 0.25 mM of each dNTP, 1x reverse transcriptase primer, 49.95 U MultiScribe^{TM} reverse transcriptase enzyme, and 3.75 U RNase inhibitor in a 15- µl total reaction volume. Reactions were incubated according to the manufacturer´s instructions. The microRNA qPCR reaction contained 1.3 µl of diluted (1:2) reverse transcription reaction, 1x TaqMan assay®, and 1x TaqMan Universal Master Mix II without UNG (Thermo Fisher Scientific, Waltham, MA) in a total reaction volume of 10 µl. All microRNA qPCR reactions were done in triplicate using amplification cycle amount of 45.

Guidelines for the interpretation of microRNA amplification results or the accepted range of variation between replicates have not been established. Here, the sample was determined as positive if microRNA amplification was detected in two or more of the replicates. For positive samples, a normalized mean threshold cycle (ΔCt) for both bkv-miR-B15p and bkv-miR-B1-3p were calculated by subtracting the mean Ct value of cel-miR-39-3p from the mean Ct values of both viral microRNAs. To compare microRNA expression between symptomatic kidney transplant patients and asymptomatic controls, the mean ΔCt of both controls was first subtracted from ΔCt of each sick to obtain ΔΔCt and the fold change of microRNA expression was calculated according to the $2^{-\Delta\Delta Ct}$ equation (251).

4 RESULTS AND DISCUSSION

Mutations frequently occur within the genome of human polyomaviruses that might alter viral pathogenicity. Mutations within the JCPyV genome are prerequisite for PML onset possibly increasing the viral replication efficiency and altering the host cell tropism (7-9, 182). Similarly, mutations within the BKPyV genome might also have a role in the viral pathogenicity and development of BKPVAN (10), although only few studies on the matter currently exist. The mechanisms behind the emergence of mutated strains as well as the exact content of mutations needed for the development of PML and polyomavirus-associated nephropathy remain still elusive.

In this thesis, the role of sequence variation of JCPyV and BKPyV in the pathogenesis of PML and polyomavirus-associated nephropathy was investigated by characterizing JCPyV and BKPyV strains from three PML patients (I), twenty-one kidney transplant patients with either stable graft function or histologically confirmed JCPVAN (II), and twenty-four kidney transplant patients with or without clinical or histological evidence for BKPVAN (III, IV) using novel MPS techniques.

4.1 Characterization of complete JC polyomavirus genomes associated with PML and JCPVAN (I, II)

Replication of reactivated JCPyV in the brain glial cells may lead to destruction of white brain matter, a hallmark for PML onset (1). PML is always preceded by mutations within the JCPyV genome and subsequent development of neurotropic strains (175) but the exact content of mutations needed for the disease onset is elusive. Extremely rarely reactivation of JCPyV replication in the kidneys can result in the lytic infection of kidney tubular epithelial cells and development of JCPVAN (3, 4). Owing to the rarity of this condition, little is known about the role of viral genetic variation in JCPVAN pathogenesis.

In this thesis, the role of mutations within the JCPyV genome in the development of the rare PML and JCPVAN diseases was investigated by characterizing all mutations originating from single JCPyV genomes obtained from PML patients (I) and kidney transplant patients with either histologically confirmed JCPVAN or stable graft function (II). Cerebrospinal fluid samples from two immunosuppressed PML patients (patients 1 and 2) and from one patient without any underlying immunosuppressive medication or condition (patient 3) were collected. The course of PML was rapid in all patients and they died within months from diagnosis. Also included were urine samples of twenty stable kidney transplant patients (patients 4-23) and the urine and plasma samples of one patient (patient 24) with histologically confirmed JCPVAN. The patient with JCPVAN eventually lost the graft due to the primary JCPyV infection and antibody-mediated rejection.

In order to amplify and characterize a complete JCPyV genome, a qualitative PCR method with forward and reverse primers binding to the adjacent regions within the JCPyV LTag gene (nucleotides 3642…3663 and 3620…3641, respectively) was established. The aim was to sequence every nucleotide in the JCPyV genome using PacBio SMRT sequencing and characterize complete JCPyV strains in single sequence reads without bias of assembling smaller sequence fragments.

4.1.1 Sequence data

Sufficient amounts of sequence subreads (mean: 25,111; range: 17,260–32,537) for characterization of JCPyV strains were obtained from all PML patients due to the

high-level JCPyV replication (range: 16,178–937,900 copies/mL) in the cerebrospinal fluid (**Table 16**). More detailed information on sequence data is presented in **Supplementary Table 1**. The subreads were used to generate consensus circular sequences for all samples using the Long Amplicon Analysis tool: three different consensus circular sequences representing complete JCPyV sequences 1.1, 1.2, and 1.3 with subread coverages of 38, 159, and 142, respectively, were obtained from patient 1. The circular consensus sequences of patients 2 and 3 represented both a single JCPyV sequence with high coverage (>500 subreads). Of note, 8.7% of the consensus sequences from PML patient 1 were interpreted as low-quality noise sequences generated during sequencing.

Sequencing of complete JCPyV genomes from the urine of kidney transplant patients yielded also high amount of sequence subreads, on average 18,873 (range: 9,060–27,911). All stable kidney transplant patients had high-level JCPyV viruria with average JCPyV load 8.84E+07 copies/mL (range: 3.50E+03–7.01E+08). None of the patients had simultaneous BKPyV viremia (<400 copies/mL) or viruria (<3000 copies/mL). The circular consensus sequences of fourteen patients represented a single JCPyV sequence with high subread coverage (>500) and average length of 5,162 bp (range: 5,161–5,164); further, two distinct JCPyV sequences with an average coverage of 245.5 (range 62–438) were detected in the urine of five patients. Three patients (11, 20, 22) had on average 4.1% (range: 2.2–5.8) of circular consensus sequences interpreted as noise sequences. The patient with biopsy-confirmed JCPVAN had an average 9,860.25 subreads obtained from the urine that all represented a single JCPyV sequence. Noise sequences (5.5%) were detected in the urine sample 4. No complete JCPyV sequences were obtained from the plasma.

Table 16. Summary of sequence data obtained from PML patients, stable kidney transplant patients, and a kidney transplant patient with biopsy-confirmed JCPVAN. The average values for twenty stable kidney transplant patients are shown. For more detailed information, see Supplementary Table 1.

Study	Patient(s)	Source	JCPyV load (copies/mL)	Total amount of subreads	Strain	Subread coverage	CCS length (bp)
I	1	CSF	937,900	17,260	1.1	38	5,127
					1.2	159	5,233
					1.3	142	5,285
	2	CSF	18,365	32,537	2	>500	5,151
	3	CSF	16,178	25,537	3	>500	5,176
II	4-23	Urine	7.83E+07	18,873	4-23	>500	5,162
	24	Plasma 1	<400	1,842	N/A	N/A	N/A
		Plasma 2	<400	3,529	N/A	N/A	N/A
		Plasma 3	<400	6,180	N/A	N/A	N/A
		Urine 1	6.11E+07	5,008	24.1	>500	5,164
		Urine 2	6.50E+07	6,346	24.2	>500	5,163
		Urine 3	9.26E+06	5,875	24.3	>500	5,164
		Urine 4	8,400	22,212	24.4	>500	5,162

Abbreviations: CSF, cerebrospinal fluid; JCPyV, JC polyomavirus; bp, base pair; CCS, circular consensus sequence.

4.1.2 NCCR

JCPyV NCCR can be divided into arbitrary sequence blocks A (25 bp), B (23 bp), C (55 bp), D (66 bp), E (18 bp), and F (80 bp) (65); the blocks A to F constitute the TCR containing promoters and enhancer elements for early and late viral gene expression. Based on the architecture of TCR, JCPyV strains can be classified into archetype strains with small TCR point mutations and neurotropic strains with larger TCR rearrangements (175). The development of PML is always associated with mutations within JCPyV NCCR; such mutations may increase viral replication efficiency by increasing the number of transcription factor binding sites favouring viral replication in the brain glial cells (8). In this thesis, mutated NCCR regions were characterized from the cerebrospinal fluid of PML patients (I) (**Table 17**, **Figure 13**). The NCCR was the most variable viral gene region. The architecture of NCCR was similar, although not identical, between different JCPyV strains, and the length varied both between the strains of different patients (strain 2: 374 and strain 3: 397 bp) and between the strains of a single patient (strain 1.1: 313 bp; strain 1.2: 454 bp; strain 1.3: 506 bp). Of note, patient 1 with multiple neurotropic strains had an exceptionally rapid onset of PML. Although our cohort is limited, it seems that the content of mutations within the JCPyV genome rather than presence of multiple viral strains is important for PML pathogenesis.

For classification, TCR regions of patient strains were compared to the corresponding region of the archetype JCPyV CY reference strain (65). Similar rearrangements were detected in the strains 1.1, 1.2, and 1.3 of PML patient 1 indicating that they could be derivatives of a single NCCR region. Prolonged clinical picture of this patient may have enabled the emergence of multiple neurotropic strains. Based on the results, the deletion of archetypal D block, harboring binding sites for glial cell-specific viral transcription activators Spi-B and Sp1 (179, 252), seems to occur frequently in neurotropic strains in agreement with previous studies (63, 113, 176-178). Deletion of D block seems not necessary for PML onset, however, as no such mutation was observed in the neurotropic strain of PML patient 2. Moreover, either partial or complete duplications seem to take place in the archetypal A, B, C, E, and F blocks of neurotropic strains; these blocks harbor also binding sites for Spi-B and Sp1 indicating that the loss of binding sites in one block could be complemented by an increase in the number of binding sites in other blocks. No direct correlation between a specific NCCR mutation and clinical picture was observed. Interestingly, the patient 2 with rapid clinical course had only an archetype-like strain with 10-bp deletion in the archetypal F block (173, 253) suggesting that even minute NCCR modifications may be sufficient for PML onset and unknown patient-specific factors could play a role in the disease pathogenesis. Finally, single nucleotide point mutations were detected both in the F block (G217A, 220delG) and outside the TCR adjacent to the early gene region (T5011A, T5020C, and G5033C) but the putative biological function of these mutations remains to be determined.

Table 17. JCPyV NCCR regions obtained from the cerebrospinal fluid of PML patients. Parenthesis indicate a truncated archetypal sequence block.

Patient	Strain	NCCR (bp)	TCR (bp)	Architecture	Phenotype	Point mutation(s)
1	1.1	313	185	A–B–(C)–E–F	Neurotropic	G217A, 220delG, T5011A, T5020C, G5033C
	1.2	454	326	A–B–(C)–E–F–(C)–E–F	Neurotropic	G217A, 220delG, T5011A, T5020C, G5033C
	1.3	506	378	A–B–(C)–E–(F)–A–B–(C)–E–F	Neurotropic	G217A, 220delG, T5011A, T5020C, G5033C
2	2	374	246	A–B–C–D–E–(F)	Archetype-like	G217A, T5011A, T5020C, G5033C
3	3	397	269	A–B–(C)–E–(B)–(C)–E–F	Neurotropic	G217A

Abbreviations: NCCR, noncoding control region; TCR, transcriptional control region.

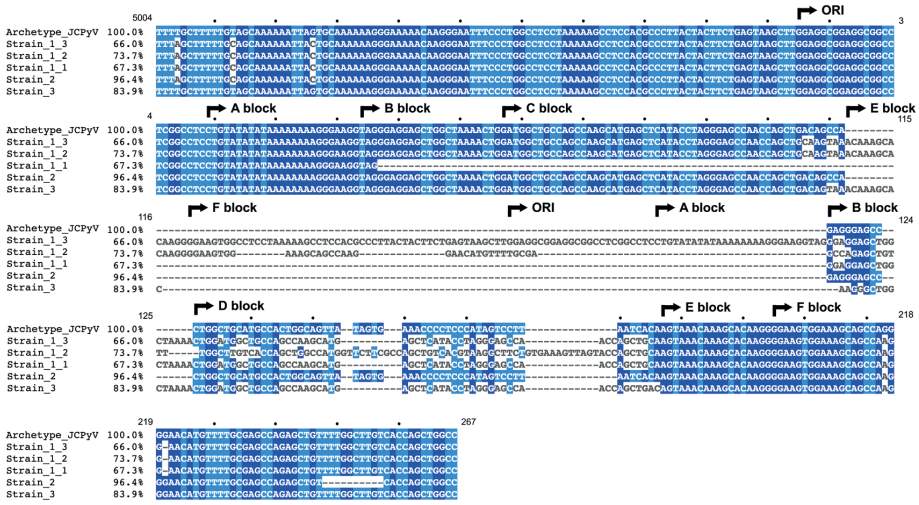


Figure 13. JCPyV NCCR regions obtained from the cerebrospinal fluid of three PML patients. JCPyV CY strain (AB038249) was used as an archetype reference strain. For multiple sequence alignment, the multiple sequence alignment tool Kalign and multiple sequence alignment viewer MView (EMBL-EBI) with default parameters were used.

NCCR rearrangements of neurotropic and archetype-like JCPyV strains of PML patients affected the number of binding sites for host cell transcription factors known to regulate JCPyV replication and gene expression (**Table 18**; Liimatainen H. et al., unpublished results). First, the duplications of archetypal A, B, C, and F blocks increased the binding sites for NF-1, AP-1, p53, and Tst-1. These factors have been shown to increase viral replication in the susceptible host cells: Tst-1 increases viral early and late gene expression, and AP-1 regulates viral replication by binding to LTag (254, 255). Members of NF-1 transcription factor family, NF-1X and NF-1D, have been shown to increase JCPyV replication in glial cells and hematopoietic stem cells (180,

256, 257). Further, most variation between the strains was observed in the number of binding sites for Spi-B known to increase JCPyV replication in glial cells (179); the amount of Spi-B binding sites within the TCR was either increased (strains 1.2, 1.3, and 2) or decreased (strains 1.1 and 3).

Table 18. Potential transcription factor binding sites in the TCR of PML-associated JCPyV strains. The number of binding sites for each transcription factor within the TCR of archetype JCPyV CY reference strain (AB038249) is shown in parenthesis.

Strain	Architecture	NF-1 (2)	Sp1 (2)	Tst-1 (1)	AP-1 (2)	Egr-1 (0)	Spi-B (5)	p53 (1)
1.1	A–B–(C)–E–F	2	1	1	2	0	4	1
1.2	A–B–(C)–E–F–(C)–E–F	3	1	1	3	0	7	2
1.3	A–B–(C)–E–(F)–A–B–(C)–E–F	3	2	2	3	1	6	1
2	A–B–C–D–E–(F)	2	2	1	2	0	5	1
3	A–B–(C)–E–(B)–(C)–E–F	3	1	1	3	0	4	1

Abbreviations: NF-1, nuclear factor 1; Sp1, specificity protein 1; Tst-1, POU domain transcription factor; AP-1, activator protein 1; Egr-1, early growth response protein 1; Spi-B, hematopoietic transcription factor; p53, tumor suppressor protein p53.

Partly due to the rarity of JCPVAN (3, 4), it remains unknown whether mutations within the JCPyV genome are associated with the development of JCPVAN. In the present work, archetype JCPyV was found dominant both in the urine of twenty kidney transplant patients with stable graft function and in the patient with biopsy-confirmed JCPVAN (II) (**Tables 19** and **20**, **Supplementary Figure 1**). Compared with archetype reference CY strain, the patient strains shared 99% nucleotide sequence identity with the reference archetype strain. Archetype JCPyV is thought to cause asymptomatic primary infections already in the childhood (65, 172-174). Similar to PML patient strains, all archetype strains had a point mutation G217A within the F block; additional point mutations G95C, G108A, A234T, and T229C were observed in the TCR of strains from four stable kidney transplant patients (21, 6, 18, and 19, respectively). Of note, patient 14 with minor graft dysfunction had an archetype-like strain harboring 8-bp and 6-bp deletions in the archetypal B and C blocks, respectively, deleting one binding sites for NF-1 known to increase viral replication in glial cells (180, 256, 257); whether NF-1 increases the replication of JCPyV in the kidneys remains to be studied. Similar deletion has been previously detected in JCPyV strains of stable kidney transplant patients (174).

Table 19. JCPyV strains from urine and plasma of a kidney transplant patient with biopsy-confirmed JCPVAN.

Sample	Strain	NCCR (bp)	TCR (bp)	Architecture	Phenotype	Point mutations
Plasma	N/A	N/A	N/A	N/A	N/A	N/A
Plasma	N/A	N/A	N/A	N/A	N/A	N/A
Plasma	N/A	N/A	N/A	N/A	N/A	N/A
Urine	24.1	384	256	A–B–C–D–E–F	Archetype	G217A
Urine	24.2	384	256	A–B–C–D–E–F	Archetype	G217A
Urine	24.3	384	256	A–B–C–D–E–F	Archetype	G217A
Urine	24.4	384	256	A–B–C–D–E–F	Archetype	G217A

Abbreviations: NCCR, noncoding control region; TCR, transcriptional control region; bp, base pair; N/A, not available.

Table 20. JCPyV strains from urine of twenty kidney transplant patients with stable graft function. Patient 14 had minor graft dysfunction and patient 20 experienced pollakisuria and pyuria. Truncated archetypal sequence blocks are presented in parenthesis.

Patient	Strain	NCCR (bp)	TCR (bp)	Architecture	Phenotype	Point mutations
4	4	384	256	A–B–C–D–E–F	Archetype	G217A
5	5	384	256	A–B–C–D–E–F	Archetype	G217A
6	6	384	256	A–B–C–D–E–F	Archetype	G108A, G217A
7	7	384	256	A–B–C–D–E–F	Archetype	G217A
8	8	384	256	A–B–C–D–E–F	Archetype	G217A
9	9	384	256	A–B–C–D–E–F	Archetype	G217A
10	10	384	256	A–B–C–D–E–F	Archetype	G217A
11	11	384	256	A–B–C–D–E–F	Archetype	G217A
12	12.1	384	256	A–B–C–D–E–F	Archetype	G217A
	12.2	384	256	A–B–C–D–E–F	Archetype	G217A
13	13	384	256	A–B–C–D–E–F	Archetype	G217A
14	14	370	242	A–(B)–(C)–D–E–F	Archetype-like	G217A, T229C
15	15	384	256	A–B–C–D–E–F	Archetype	G217A
16	16.1	384	256	A–B–C–D–E–F	Archetype	G217A
	16.2	384	256	A–B–C–D–E–F	Archetype	G217A
17	17	384	256	A–B–C–D–E–F	Archetype	G217A
18	18.1	384	256	A–B–C–D–E–F	Archetype	G217A, A234T
	18.2	384	256	A–B–C–D–E–F	Archetype	G217A
19	19	384	256	A–B–C–D–E–F	Archetype	G217A, T229C
20	20	384	256	A–B–C–D–E–F	Archetype	G217A
21	21.1	384	256	A–B–C–D–E–F	Archetype	G95C, G217A
	21.2	384	256	A–B–C–D–E–F	Archetype	G217A
22	22.1	384	256	A–B–C–D–E–F	Archetype	G217A
	22.2	384	256	A–B–C–D–E–F	Archetype	G217A
23	23	384	256	A–B–C–D–E–F	Archetype	G217A

Abbreviations: NCCR, noncoding control region; TCR, transcriptional control region; bp, base pair.

4.1.3 Genotypes

Genotyping was based on the comparison of nucleotide sequence of the protein coding region of patient strains to the corresponding region of the reference genotype strains (**Supplementary Table 2**) (115). A total of seven JCPyV genotypes (1-4, 6-8) are currently known, and genotypes 1, 2, and 3 are further divided into several genotypes (110). All study patients had only one JCPyV genotype. Strains of two PML patients and seventeen stable kidney transplant patients with JCPyV viremia belonged to genotype 1: subtypes 1A and 1B were detected in nine and ten patients, respectively. Results further support the notion that genotype 1 is frequently detected in Europe (114, 115). The strain of PML patient 3 shared the highest nucleotide identity with the Asian genotype strain 2B (114, 115). Although JCPyV genotype 2 has been shown to more often associate with PML (113), neither clinical characteristics nor outcome of this patient differed from other PML patients. Of note, a strain of stable kidney transplant patient 8 belonged also to the genotype 2B. The genotype 4, dominant in USA and Europe (112, 114, 115), was only detected in two stable kidney transplant patients (10 and 11) and in the patient with biopsy-confirmed JCPVAN suggesting

that genotype 4 is rare in Finnish population.

Out of five stable kidney transplant patients with two distinct JCPyV strains in the urine, the differences between strains were observed in the protein coding region in four patients (12, 16, 18, and 22); the strains of patient 21 differed only by a single point mutation (G95C) within the TCR. The presence of multiple strains within an individual could be a result of co-infection of different strains as suggested for other human polyomaviruses (22, 24). The strain-specific differences were nucleotide point mutations except in patient 12 whose strains differed by a larger, 12-bp deletion in the agnoprotein (nucleotides 127…139). Similar but larger deletion (nucleotides 51…216) has been observed in the neurotropic strains of PML patients suggesting a role in the regulation of viral replication in the susceptible host cells (184).

4.1.4 Mutations within the VP1

Mutations leading to amino acid changes within VP1 were characterized in detail as mutations in the receptor recognition region of VP1 may enhance the ability of neurotropic strains to infect brain glial cells (**Table 21**) (7, 9, 182). The protein coding regions of patient strains from both PML and kidney transplant patients shared 99% nucleotide identity with the corresponding genotype strain (I, II). Mutations within the VP1 were more frequent in the archetype and archetype-like strains than in the neurotropic strains which are defined by rearrangements within the NCCR. All strains from PML patients had amino acid point mutations located in the HI loop, a part of receptor recognition region (47), indicating that alterations in the receptor binding region are necessary for the PML onset. A different HI loop mutation was observed in the three neurotropic strains of patient 1 (S267L) than in the archetype-like strain of patient 2 and the neurotropic strain of patient 3 (S269F): mutations in these amino acids might change host cell tropism towards glial cells (7, 182). The strains of patients 2 and 3 had additional mutations in the EF (L158V) and BC (K60N, E69Q) loops, respectively. The strain of patient 2 had also point mutations G8A and K345R not affecting the receptor recognition region of VP1. The neurotropic and archetype-like strains of PML patients had mutations affecting the HI and BC loops previously detected in the neurotropic strains of PML patients (7, 182).

The strains of stable kidney transplant patients had amino acid point mutations within VP1 affecting especially the EF (L158V, A158V, V195A), BC (D52N, V64I, K75R), and DE (T128S) loops; most frequent mutations V64I and L158V were found in ten and five stable kidney transplant patients, respectively. The archetype strains of four stable kidney transplant patients and the archetype strain of patient 14 with minor graft dysfunction had the same point mutations G8A, L158V, and K345R as seen in the neurotropic strain of patient 2. Most VP1 mutations, such as S222N detected in strains of ten stable kidney transplant patients, were, however, not located in the receptor recognition region of VP1. Further, the strain of patient 8 had no VP1 amino acid point mutations. The results suggest that mutations within the VP1 occasionally occur during viral replication in the kidneys and brain but different mutations may be favored due to different biology of these cells.

The archetype strains of the patient with JCPVAN had two mutations affecting the EF loop (A158V, V195A) and two amino acid point mutations not located in the VP1 loops (V86I, E335G). The same point mutations were identified also in two stable kidney transplant patients but none of the neurotropic strains of PML patients, now or previously (7, 182). These results show that mutations in the receptor recognition region of VP1 occur more frequently in the strains associated with PML than JCPVAN: especially mutations within the HI loop seem to have a role in the host cell tropism.

Table 21. Amino acid point mutations within the VP1 of JCPyV strains of PML patients and kidney transplant patients with either stable graft function or histologically confirmed JCPVAN.

Patient	Status	Strain	Point mutation(s)		VP1 loop affected
			Nucleic acid	Amino acid	
1	PML	1.1	C800T	S267L	HI
		1.2	C800T	S267L	HI
		1.3	C800T	S267L	HI
2	PML	2	G23C	G8A	None
			C472G, T474A	L158V	EF
			C806T	S269F	HI
			A1034G	K345R	None
3	PML	3	G180C	K60N	BC
			G205C	E69Q	BC
			C806T	S269F	HI
4	KTx	4	G23C	G8A	None
			C472G	L158V	EF
			A1034G	K345R	None
5	KTx	5	G23C	G8A	None
			C472G	L158V	EF
			A1034G	K345R	None
6	KTx	6	A224G	K75R	BC
7	KTx	7	G190A	V64I	BC
			G665A	S222N	None
8	KTx	8	-	-	-
9	KTx	9	G23C	G8A	None
			C472G	L158V	EF
			A1034G	K345R	None
10	KTx	10	G256A	V86I	None
			C473T	A158V	EF
			T584C	V195A	None
			A1004G	E335G	None
11	KTx	11	G256A	V86I	None
			C473T	A158V	EF
			T584C	V195A	None
			A1004G	E335G	None
12	KTx	12.1	G190A	V64I	BC
			G665A	S222N	None
		12.2	G190A	V64I	BC
			G665A	S222N	None
13	KTx	13	G190A	V64I	BC
			G665A	S222N	None
14	KTx	14	G23C	G8A	None
			C472G	L158V	EF
			A1034G	K345R	None
15	KTx	15	G190A	V64I	BC
			G665A	S222N	None

Table 21. continued.

Patient	Status	Strain	Point mutation(s)		VP1 loop affected
			Nucleic acid	Amino acid	
16	KTx	16.1	G190A	V64I	BC
			G665A	S222N	None
		16.2	G190A	V64I	BC
			G665A	S222N	None
17	KTx	17	G190A	V64I	BC
			G665A	S222N	None
18	KTx	18.1	C383G	T128S	DE
		18.2	C383G	T128S	DE
19	KTx	19	G23C	G8A	None
			G154A	D52N	BC
			C472G	L158V	EF
			A1034G	K345R	None
20	KTx	20	G190A	V64I	BC
			G665A	S222N	None
21	KTx	21.1	G190A	V64I	BC
			G665A	S222N	None
		21.2	G190A	V64I	BC
			G665A	S222N	None
22	KTx	22.1	G190A	V64I	BC
			G665A	S222N	None
		22.2	G154A	D52N	BC
			G190A	V64I	BC
			G665A	S222N	None
23	KTx	23	G190A	V64I	BC
			G665A	S222N	None
24	JCPVAN	24.1-24.4	G256A	V86I	None
			C473T, T474G	A158V	EF
			T584C	V195A	EF
			A1004G	E335G	None

Abbreviations: PML, progressive multifocal leukoencephalopathy; KTx, kidney transplantation; JCPVAN, JC polyomavirus-associated nephropathy; VP1, viral protein 1.

4.1.5 Mutations within the LTag

As a major regulatory protein, LTag is a key regulator of viral DNA replication and gene expression (77, 78). Thus, similar to VP1 and NCCR, mutations within the LTag might alter viral pathogenicity and have a role in the development of PML and JCPVAN. To this end, amino acid point mutations within the LTag of JCPyV strains of PML patients and kidney transplant patients with either stable graft function or biopsy-confirmed JCPVAN were investigated (II, Liimatainen H. et al., unpublished results) (**Table 22**). Of the PML patients, only the neurotropic strains of patient 1 and the archetype-like strain of patient 2 had mutations within the LTag (K507R and S653F, respectively). Further, only the mutation K507R affected the functional domain (helicase) of LTag that might inhibit viral replication efficiency (183). No mutations were detected in the immunocompetent PML patient 3.

Mutations located in the origin binding domain of LTag (L160F, A161S) were most frequent in stable kidney transplant patients. A zinc finger domain mutation T357S and the helicase domain mutation K507R were observed in patients 17 and 18, respectively. Finally, a total of eight stable kidney transplant patients and the patient with JCPVAN did not have mutations affecting the functional domains of LTag. Mutations in the functional domains of LTag seem to rarely associate with the development of PML and never with JCPVAN indicating a different role of LTag in different host cells; further, the patient-specific factors of PML patients, such as prolonged immunosuppression, might allow accumulation of mutations within the LTag.

Table 22. Amino acid point mutations within the LTag of JCPyV strains of PML patients and kidney transplant patients with either stable graft function or histologically confirmed JCPVAN.

Patient	Status	Strain	Point mutation(s)		LTag domain affected
			Nucleic acid	Amino acid	
1	PML	1.1	A1520G	K507R	Helicase
		1.2	A1520G	K507R	Helicase
		1.3	A1520G	K507R	Helicase
2	PML	2	C1958T	S653F	None
3	PML	3	-	-	-
4	KTx	4	C1958T	S653F	None
5	KTx	5	C1958T	S653F	None
6	KTx	6	C1958T	S653F	None
7	KTx	7	C478T	L160F	OBD
8	KTx	8	G481T	A161S	OBD
9	KTx	9	C1958T	S653F	None
10	KTx	10	G1820A	G607E	None
11	KTx	11	G1820A	G607E	None
12	KTx	12.1	C478T	L160F	OBD
		12.2	C478T	L160F	OBD
13	KTx	13	C478T	L160F	OBD
			C1958T	S653F	None
14	KTx	14	C1958T	S653F	None
15	KTx	15	C478T	L160F	OBD
16	KTx	16.1	C478T	L160F	OBD
		16.2	C478T	L160F	OBD
17	KTx	17	C478T	L160F	OBD
			A1069T	T357S	Zinc finger
18	KTx	18.1	A1520G	K507R	Helicase
		18.2	A1520G	K507R	Helicase
19	KTx	19	C1958T	S653F	None
20	KTx	20	C478T	L160F	OBD
			C1958T	S653F	None

Table 22. continued.

Patient	Status	Strain	Point mutation(s)		LTag domain affected
			Nucleic acid	Amino acid	
21	KTx	21.1	C478T	L160F	OBD
		21.2	C478T	L160F	OBD
22	KTx	22.1	C478T	L160F	OBD
			C1958T	S653F	None
		22.2	C478T	L160F	OBD
			C1958T	S653F	None
23	KTx	23	C478T	L160F	OBD
			C1958T	S653F	None
24	JCPVAN	24.1-24.4	G1820A	G607E	None

Abbreviations: PML, progressive multifocal leukoencephalopathy; KTx, kidney transplantation; JCPVAN, JC polyomavirus-associated nephropathy; LTag, large T antigen. OBD, origin binding domain.

4.2 BKPyV TCR rearrangements in BKPVAN (III, IV)

Deterioration of the immune system especially after kidney transplantation may predispose to reactivation of BKPyV replication in the kidneys and lead to the development of BKPVAN. BKPVAN is a major cause of allograft dysfunction and rejection in kidney transplant patients (186, 188, 189). Similar to JCPyV, rearrangements, deletions, and duplications frequently occur within the BKPyV TCR that might increase viral replication efficiency by altering the number of binding sites for host cell transcription factors (10). Although BKPyV strains with rearranged TCR regions are detected in kidney transplant patients with and without BKPVAN (10, 122, 206), whether TCR mutations are associated with the pathogenesis of BKPVAN remains still unknown.

In this thesis, the role of mutations within the BKPyV TCR in the development of BKPVAN was studied by characterizing BKPyV TCR regions from a total of twenty-four kidney transplant patients with varying kidney function (III, IV). Two patients had no clinical evidence for BKPVAN progression accompanied with low-level (<4 log10 copies/mL) BKPyV viremia (probable BKPVAN). A total of ten patients had high-level (>4 log10 copies/mL) BKPyV viremia (presumptive BKPVAN) and twelve patients had BKPVAN confirmed histologically (definitive BKPVAN). Plasma samples from a total of seven definitive BKPVAN patients and two presumptive BKPVAN patients (III) were collected while pairs of urine and plasma samples were obtained from two probable BKPVAN patients, eight presumptive BKPVAN patients, and five definitive BKPVAN patients (IV). To characterize BKPyV strains associated with the development of BKPVAN, a PCR method for amplification of BKPyV TCR was established. The aim was to sequence the complete BKPyV TCR in a single continuous read using Illumina short-read sequencing.

4.2.1 Sequence data

Paired-end sequencing of complete BKPyV TCR from plasma of nine kidney transplant patients (III) and pairs of urine and plasma of fifteen kidney transplant patients (IV) yielded sufficient amount of sequences for the characterization of BKPyV TCR (**Supplementary Table 3**). The original read amount obtained from a single sample was on average 216,616 read-pairs (range: 143,719–360,608). After subjecting the read-pairs to the in-house pipeline, on average 159,513 sequences (range: 85,959–

263,356) representing on average 73.4% (range: 57.3–88.6) of all read-pairs in a sample were available for the characterization of BKPyV TCR.

4.2.2 TCR

Similar to JCPyV, BKPyV strains can be classified into two groups based on the architecture of TCR: archetype strains having small point mutations within the TCR and mutated strains with highly variable TCR structure (203). Asymptomatic primary infections are thought to be caused by archetype strains while rearrangements within the TCR might precede the development of BKPVAN (10, 172). To this end, BKPyV TCR regions from plasma of twenty-four kidney transplant patients with either probable, presumptive, or definitive BKPVAN (III, IV) were characterized. All patients had BKPyV viremia and the BKPyV DNA load in the plasma was on average 1.90E+07 copies/mL (range: 2.0E+03–4.20E+08).

In twenty-three patients, majority (on average 91.6%) of all sequences in the plasma represented archetype TCR; in the plasma of six presumptive BKPVAN and two definitive BKPVAN patients only archetype BKPyV TCR was characterized (**Table 23**). The nucleotide identity of archetype TCR regions was 97.01–100% with the TCR of archetype BKPyV WW reference strain (203): a total of 26 nucleotide point mutations were identified (**Supplementary Figure 2**). Of these, ten mutations (T65C, C179G, G181A, G182T, T212C, C239G, C326G, G328A, A369G, T373G) were located in the binding sites for transcription factor NF-1 and two mutations (del88_89, A90C) were located in the binding sites for tumor suppressor protein p53; both NF-1 and p53 are known to increase BKPyV replication (71, 258). Similar archetype-like strains with minute TCR modifications are occasionally detected in kidney transplant patients with and without BKPVAN (10). Whether these mutations decrease the replication efficiency of rearranged strains needs to be confirmed.

Intriguingly, a large proportion of sequences (46.93%) in the plasma of one definitive BKPVAN patient (patient 20) represented a rearranged TCR with 18-bp-deletion in the archetypal P block; a small proportion (2.60%) of sequences represented archetype TCR. Rearranged BKPyV TCR regions were also characterized from the plasma of two probable BKPVAN patients, three presumptive BKPVAN patients, and ten definitive BKPVAN patients (**Table 23**). Rearranged TCR regions were especially abundant in the definitive BKPVAN patients: the average proportions of rearranged TCR regions in the probable BKPVAN patients, presumptive BKPVAN patients, and definitive BKPVAN patients were 2.96%, 1.49%, and 13.94%, respectively. Statistically significant difference was not reached ($p>0.05$), however, probably due to the small cohort size. Most variation in the structure of rearranged TCR regions was observed in the definitive BKPVAN patients: while these patients had 1 to 15 different mutated strains in the plasma, a maximum of two mutated strains were identified in other patient groups. Increased viral replication in the kidneys due to BKPVAN might allow accumulation of mutations within the archetypal TCR and subsequent development of mutated strains. Most frequent TCR mutation in all three patient groups was a partial deletion of archetypal Q and R blocks harbouring binding sites for transcription factors Sp1 and NF-1 that increase viral early gene expression (71, 82). Further, partial and complete duplications of P and Q blocks containing binding sites for viral transcription activators AP-1, NFAT, NF-1, Sp1, and p53 (71, 82, 258, 259), were seen in the plasma of two presumptive and six definitive BKPVAN patients. According to the results, deletions and duplications in the archetypal P, Q, and R blocks frequently occur within the TCR of mutated strains that might have a role in the development of BKPVAN.

Table 23. BKPyV TCR regions in plasma of kidney transplant patients with or without histologically confirmed BKPVAN. In case of patients 20-24, only TCR regions representing >2.6% of all sequences in a sample are shown. Nine patients (5, 7, 16, 17, 20-24) experienced graft dysfunction at sampling.

Patient	Kidney status	Strain	TCR (bp)	TCR architecture	Phenotype	Proportion (%)
10	Probable BKPVAN	10.1	234	P-Q-R-S	Archetype	98.17
		10.2	207	P-(Q)-(R)-S	Rearranged	1.93
11	Probable BKPVAN	11.1	233	P-Q-R-S	Archetype	96.01
		11.2	188	(P)-(Q)-R-S	Rearranged	3.98
3	Presumptive BKPVAN	3	233	P-Q-R-S	Archetype	99.90
9	Presumptive BKPVAN	9.1	234	P-Q-R-S	Archetype	97.60
		9.2	315	P-Q-(P)-Q-R-S	Rearranged	2.10
12	Presumptive BKPVAN	12	233	P-Q-R-S	Archetype	99.97
13	Presumptive BKPVAN	13.1	233	P-Q-R-S	Archetype	98.90
		13.2	279	P-(Q)-(P)-Q-R-S	Rearranged	0.97
14	Presumptive BKPVAN	14	233	P-Q-R-S	Archetype	99.97
15	Presumptive BKPVAN	15	233	P-Q-R-S	Archetype	99.97
16	Presumptive BKPVAN	16	233	P-Q-R-S	Archetype	99.81
17	Presumptive BKPVAN	17	234	P-Q-R-S	Archetype	98.84
18	Presumptive BKPVAN	18	233	P-Q-R-S	Archetype	98.82
19	Presumptive BKPVAN	19.1	233	P-Q-R-S	Archetype	95.16
		19.2	184	P-(Q)-(R)-S	Rearranged	0.89
		19.3	215	P-(Q)-(R)-S	Rearranged	0.51
1	Definitive BKPVAN	1.1	233	P-Q-R-S	Archetype	96.50
		1.2	331	P-(Q)-(P)-Q-R-S	Rearranged	3.0
2	Definitive BKPVAN	2.1	233	P-Q-R-S	Archetype	97.20
		2.2	175	P-(Q)-(R)-S	Rearranged	2.70
4	Definitive BKPVAN	4.1	234	P-Q-R-S	Archetype	96.50
		4.2	275	P-(Q)-(P)-Q-R-S	Rearranged	2.50
5	Definitive BKPVAN	5.1	234	P-Q-R-S	Archetype	89.80
		5.2	180	P-Q-(R)-S	Rearranged	5.90
		5.3	283	P-(Q)-(P)-Q-R-S	Rearranged	1.20
		5.4	317	P-Q-(P)-Q-R-S	Rearranged	1.20
6	Definitive BKPVAN	6.1	234	P-Q-R-S	Archetype	83.90
		6.2	283	P-(Q)-(P)-Q-R-S	Rearranged	15.90
7	Definitive BKPVAN	7.1	234	P-Q-R-S	Archetype	77.60
		7.2	192	P-(Q)-(R)-S	Rearranged	22.40
8	Definitive BKPVAN	8	234	P-Q-R-S	Archetype	99.90

Table 23. continued.

Patient	Kidney status	Strain	TCR (bp)	TCR architecture	Phenotype	Proportion (%)
20	Definitive BKPVAN	20.1	216	(P)-Q-R-S	Rearranged	46.93
		20.2	266	P-(P)-Q-R-S	Rearranged	7.19
		20.3	266	P-(Q)-(P)-(Q)-(R)-S	Rearranged	7.14
		20.4	208	P-(Q)-(R)-S	Rearranged	5.19
		20.5	232	P-(Q)-(R)-(S)-(Q)-(R)-S	Rearranged	2.88
		20.6	225	P-(Q)-(R)-(Q)-(R)-S	Rearranged	2.60
		20.7	234	P-Q-R-S	Archetype	2.60
21	Definitive BKPVAN	21.1	234	P-Q-R-S	Archetype	57.11
		21.2	185	P-(Q)-(R)-S	Rearranged	4.75
		21.3	208	P-(Q)-(R)-S	Rearranged	4.38
22	Definitive BKPVAN	22.1	233	P-Q-R-S	Archetype	46.68
		22.2	141	P-(Q)-(S)	Rearranged	3.32
		22.3	152	P-(Q)-(S)	Rearranged	3.12
		22.4	206	P-(Q)-(R)-S	Rearranged	2.85
23	Definitive BKPVAN	23.1	233	P-Q-R-S	Archetype	46.27
		23.2	184	P-(Q)-(R)-S	Rearranged	12.86
		23.3	226	P-(Q)-(R)-(P)-(Q)-(R)-S	Rearranged	10.81
		23.4	173	P-(Q)-(R)-S	Rearranged	5.41
		23.5	145	P-(P)-(Q)-(S)	Rearranged	5.19
		23.6	215	(P)-Q-R-S	Rearranged	4.28
		23.7	113	P-(Q)-(S)	Rearranged	2.60
24	Definitive BKPVAN	24	234	P-Q-R-S	Archetype	99.42

Abbreviations: BKPVAN, BK polyomavirus-associated nephropathy; BKPyV, BK polyomavirus; TCR, transcriptional control region.

From fifteen kidney transplant patients with BKPyV viremia, urine samples were also available for characterization of BKPyV TCR regions (IV). All patients had high-level BKPyV viruria with an average BKPyV DNA load of 5.10E+08 copies/mL in the urine (range: 3.10E+07–2.0E+11). Similar to plasma, a majority of sequences represented archetype BKPyV TCR (**Table 24**). The archetype TCR regions from the urine and plasma of a single patient were always 100% identical indicating that robust viral replication in the kidneys might enable spreading of BKPyV strains from the kidneys to the systemic circulation. In two patients with either probable or definitive BKPVAN, archetype TCR predominated in the urine contrary to plasma where multiple rearranged strains were detected. In one presumptive BKPVAN patient a rearranged TCR region was identified in the urine but not in the plasma. Although rearranged TCR regions were not identical in the plasma and urine of a single patient, similar deletions of archetypal Q and R blocks were found.

Table 24. BKPyV TCR regions in urine of kidney transplant patients with or without histologically confirmed BKPVAN. In case of patients 20-24, only TCR regions representing >2.6% of all sequences in a sample are shown. Seven patients (16, 17, 20-24) experienced graft dysfunction at sampling.

Patient	BKPVAN	Strain	TCR (bp)	TCR architecture	Phenotype	Proportion (%)
10	Probable BKPVAN	10.1	234	P-Q-R-S	Archetype	97.19
		10.2	184	P-(Q)-(R)-S	Rearranged	1.76
11	Probable BKPVAN	11	233	P-Q-R-S	Archetype	99.72
12	Presumptive BKPVAN	12	233	P-Q-R-S	Archetype	99.87
13	Presumptive BKPVAN	13.1	233	P-Q-R-S	Archetype	98.72
		13.2	144	P-(Q)-(S)	Rearranged	0.74
14	Presumptive BKPVAN	14	233	P-Q-R-S	Archetype	98.1
15	Presumptive BKPVAN	15	233	P-Q-R-S	Archetype	98.76
16	Presumptive BKPVAN	16	233	P-Q-R-S	Archetype	99.26
17	Presumptive BKPVAN	17.1	234	P-Q-R-S	Archetype	92.37
		17.2	185	P-(Q)-(R)-S	Rearranged	3.88
18	Presumptive BKPVAN	18	233	P-Q-R-S	Archetype	98.82
19	Presumptive BKPVAN	19.1	233	P-Q-R-S	Archetype	95.16
		19.2	184	P-(Q)-(R)-S	Rearranged	0.89
		19.3	215	P-(Q)-(R)-S	Rearranged	0.51
20	Definitive BKPVAN	20.1	216	(P)-Q-R-S	Rearranged	27.6
		20.2	266	P-(P)-Q-R-S	Rearranged	18.17
		20.3	226	(P)-Q-R-S	Rearranged	15.18
		20.4	234	P-Q-R-S	Archetype	8.05
		20.5	185	P-(Q)-(R)-S	Rearranged	7.67
		20.6	256	P-(P)-Q-R-S	Rearranged	6.67
		20.7	266	P-(Q)-(P)-(Q)-(R)-S	Rearranged	5.52
21	Definitive BKPVAN	21.1	234	P-Q-R-S	Archetype	73.22
		21.2	185	P-(Q)-(R)-S	Rearranged	6.48
		21.3	208	P-(Q)-(R)-S	Rearranged	4.58
22	Definitive BKPVAN	22	233	P-Q-R-S	Archetype	98.77
23	Definitive BKPVAN	23.1	233	P-Q-R-S	Archetype	34.26
		23.2	215	(P)-Q-R-S	Rearranged	17.33
		23.3	184	P-(Q)-(S)	Rearranged	7.97
		23.4	113	P-(Q)-(R)-S	Rearranged	6.10
		23.5	145	P-(P)-(Q)-(S)	Rearranged	5.18
		23.6	226	P-(Q)-(R)-(P)-(Q)-(R)-S	Rearranged	4.78
		23.7	173	P-(Q)-(R)-S	Rearranged	3.98
24	Definitive BKPVAN	24	234	P-Q-R-S	Archetype	99.71

Abbreviations: BKPVAN, BK polyomavirus-associated nephropathy; BKPyV, BK polyomavirus; TCR, transcriptional control region.

4.2.3 Effect of TCR rearrangements on transcription factor binding sites

The number of binding sites for host cell transcription factors within the TCR determines the tissue-specific gene expression of BKPyV. The rearrangements within the TCR of kidney transplant patients with probable, presumptive, or definitive BKPVAN altered the number of transcription factor binding sites for Sp1, NF-1, NFAT, AP-1, and p53, known to bind BKPyV TCR (**Table 25**) (67, 204), suggesting their importance in viral DNA replication and gene expression. The amount of binding sites for NF-1 and Sp1, established BKPyV replication-enhancing transcription factors (71, 82), was decreased mostly owing to the deletions affecting the archetypal Q and R blocks. However, no correlation in the viral load and the amount of binding sites for NF-1 or Sp1 within the rearranged TCR regions was observed further indicating the predominance of archetype strains in the urine and plasma. The number of binding sites for AP-1 and p53, known to activate BKPyV early gene expression (258, 259), remained unchanged in most patients. No clear association between the number of transcription factor binding sites and disease status was established.

Table 25. The number of transcription factor binding sites within the rearranged BKPyV TCR. For comparison, the number of binding sites for each transcription factor in the TCR of archetype BKPyV WW reference strain (AB211371.1) is shown in parenthesis. In case of patients 5, 19-23, the average values are given.

Patient	Kidney status	Material	BKPyV load (copies/mL)	AP-1 (1)	NFAT (1)	NF-1 (6)	Sp1 (2)	p53 (2)
13	Probable BKPVAN	Urine	3.70E+07	1	2	5	1	1
		Plasma	1.00E+04	1	3	6	1	1
14	Probable BKPVAN	Plasma	1.10E+04	0	2	4	2	1
9	Presumptive BKPVAN	Plasma	1.93E+04	2	2	8	2	3
11	Presumptive BKPVAN	Urine	3.10E+08	1	1	4	1	1
		Plasma	2.00E+03	2	4	6	2	3
17	Presumptive BKPVAN	Urine	2.80E+08	1	3	4	1	2
19	Presumptive BKPVAN	Urine	1.00E+10	1	3	5	1.50	2
		Plasma	1.10E+05	1	3	5.50	1	2
1	Definitive BKPVAN	Plasma	4.53E+04	2	2	8	3	3
2	Definitive BKPVAN	Plasma	1.90E+05	1	1	4	1	2
4	Definitive BKPVAN	Plasma	8.91E+04	2	1	7	2	2
5	Definitive BKPVAN	Plasma	2.17E+05	1.67	1.67	6.67	2.33	2.33
6	Definitive BKPVAN	Plasma	8.5E+03	2	2	7	2	2
7	Definitive BKPVAN	Plasma	2.96E+04	1	1	6	1	2
20	Definitive BKPVAN	Urine	1.90E+11	1.29	3.14	5.14	1.71	2.14
		Plasma	4.20E+08	1.64	2.93	5.36	1.43	2
21	Definitive BKPVAN	Urine	1.60E+11	1	2	4.38	1	1.88
		Plasma	1.10E+07	1	1.70	4.30	1	1.80
22	Definitive BKPVAN	Plasma	4.70E+05	1	1.60	4.20	1.33	1.47
23	Definitive BKPVAN	Urine	2.00E+11	1.08	2.46	5.23	1.33	1.85
		Plasma	1.70E+06	1.13	2.25	4.63	1.25	1.75

Abbreviations: AP-1, activator protein 1; NFAT, nuclear factor of activated T cells; NF-1, nuclear factor 1; Sp1, specificity protein 1; p53, tumor suppressor protein p53; BKPVAN, BK polyomavirus-associated nephropathy.

4.3 BKPyV microRNA expression in BKPVAN (III)

BKPyV encodes two microRNA molecules, bkv-miR-B1-5p and bkv-miR-B1-3p, that regulate viral early gene expression (100). Viral microRNAs seem to also help evading the host immune system and transmitting viral particles between host cells (101). Thus, BKPyV-encoded microRNAs may affect viral pathogenesis and subsequently be associated with the development of BKPVAN (108, 109). In this thesis, we studied the association of BKPyV TCR mutations and the expression of BKPyV-encoded microRNAs in nine presumptive and definitive BKPVAN patients (III). BKPyV microRNA levels from a total of 53 plasma samples were measured. Plasma samples of two adult kidney transplant patients with stable graft function and no evidence of BKPyV viremia or viruria were used as controls. The collection of more control patients was hindered by the fact that 35-57% of kidney transplant patients experience BKPyV viruria (188, 192-194).

The presumptive and definitive patients had higher BKPyV microRNA expression than controls (**Table 26**). Viral microRNA expression was also observed in the controls due to the fact that they probably had encountered BKPyV infection which now persisted in their body without viremia or viruria. The presumptive and definitive BKPVAN patients had on average 8.7-fold and 8.9-fold higher 5p and 3p microRNA expression than the controls, respectively. The presumptive BKPVAN patients had higher 3p and 5p microRNA expression than the definitive BKPVAN patients (on average 9.2-fold and 2.5-fold). The 3p microRNA expression was elevated (13.6-fold, 51.5-fold) in two definitive BKPVAN patients. The small cohort size prevented any statistical analysis. Due to the nucleotide identity of JCPyV and BKPyV 3p microRNAs (101), however, it cannot be ruled out that the measured 3p microRNA levels could have been affected by JCPyV-encoded microRNAs present in the plasma.

The expression of BKPyV microRNAs is regulated by the late promoter within the NCCR, implying that NCCR mutations might affect the expression levels of BKPyV microRNAs and subsequently viral replication (103). Low BKPyV load but high BKPyV 5p microRNA expression was observed in two patients (3, 8) with archetype BKPyV predominating in the plasma suggesting that viral 5p microRNA controls the replication of archetype strain as proposed earlier (103). In a majority of patients with both archetype and rearranged strains in the plasma, both viral microRNA expression and BKPyV DNA load were increased indicating that even minor populations of rearranged BKPyV strains may enable efficient virus replication despite simultaneous high 5p microRNA expression.

Table 26. BKPyV microRNA expression in the plasma of kidney transplant patients with presumptive or definitive BKPVAN. Patients 4 and 6 experienced a graft dysfunction at sampling. Fold changes of BKPyV microRNA expression were calculated comparing the normalized Ct values of both controls to the normalized Ct values of kidney transplant patients. Correlation between BKPyV microRNA expression and BKPyV DNA load was calculated using the Excel Correl function.

Patient	Kidney status	TCR architecture	Phenotype	Fold change of BKPyV microRNAs		Correlation of BKPyV microRNAs	
				3p	5p	3p	5p
1	Definitive BKPVAN	P-Q-R-S	Archetype	5.0	18.4	-0.19	-0.24
		P-(Q)-(P)-Q-R-S	Rearranged				
2	Definitive BKPVAN	P-Q-R-S	Archetype	13.6	19.1	-0.73	-0.84
		P-(Q)-(R)-S	Rearranged				
3	Presumptive BKPVAN	P-Q-R-S	Archetype	1.2	3.9	-1.00	1.00
4	Definitive BKPVAN	P-Q-R-S	Archetype	2.0	5.2	-0.99	-0.98
		P-(Q)-(P)-Q-R-S	Rearranged				
5	Definitive BKPVAN	P-Q-R-S	Archetype	51.5	8.8	-0.76	-0.67
		P-Q-(R)-S	Rearranged				
		P-(Q)-(P)-Q-R-S	Rearranged				
		P-Q-(P)-Q-R-S	Rearranged				
6	Definitive BKPVAN	P-Q-R-S	Archetype	1.7	9.3	-0.61	-0.26
		P-(Q)-(P)-Q-R-S	Rearranged				
7	Definitive BKPVAN	P-Q-R-S	Archetype	3.0	5.6	-0.94	0.44
		P-(Q)-(R)-S	Rearranged				
8	Definitive BKPVAN	P-Q-R-S	Archetype	0.9	2.9	1.00	1.00
9	Presumptive BKPVAN	P-Q-R-S	Archetype	1.6	5.0	-0.44	-0.39
		P-Q-(P)-Q-R-S	Rearranged				

Abbreviations: BKPVAN, BK polyomavirus-associated nephropathy; TCR, transcriptional control region.

4.4 The role of viral mutations in PML and polyomavirus-associated nephropathy

PML is a neurodegenerative brain disease with poor prognosis, and even if the disease progress could be halted the damage is irreversible resulting in various neurological disabilities. In accordance with previous studies (1, 63, 113, 176-178), our PML patients harbored neurotropic JCPyV strains in the cerebrospinal fluid suggesting that mutations within the NCCR are necessary for PML onset. In one patient multiple neurotropic strains with different NCCR structure were identified; whether distinct neurotropic strains develop before PML onset or as a result of robust viral replication in the brain remains to be determined. It seems that changes in the number of binding sites for glial cell-specific transcription factors rather than large NCCR rearrangements play a role in the pathogenesis as all PML-associated strains had changes in the number of binding sites for Spi-B, known to increase JCPyV replication in the brain glial cells (179).

Reactivation of JCPyV replication in the kidneys can extremely rarely lead to the development of JCPVAN in kidney transplant patients (3, 4). The presence of

archetype JCPyV strains in the urine of immunosuppressed and immunocompetent individuals (65, 172-174) suggests that archetype JCPyV establishes latency in the kidneys. In the present work, only archetype JCPyV was established in the urine of kidney transplant patients with stable kidney function and biopsy-confirmed JCPVAN indicating that mutations within the viral genome are not necessary for the pathogenesis of JCPVAN. Interestingly, the JCPVAN patient was previously shown to be JCPyV-seronegative prior transplantation but developed JCPyV-specific antibody response after transplantation (219). Thus, it seems that other factors, such as the lack of JCPyV-specific cell-mediated immunity as suggested for PML (135), may allow the replication of persisting archetype JCPyV in the kidneys ultimately leading to JCPVAN onset. This is in contrast with PML, where mutations within the JCPyV genome are required for the disease onset. Different roles of viral sequence variation in the pathogenesis of PML and JCPVAN could be due to different replication conditions in kidney and brain glial cells. Further, in case of JCPVAN, viral mutations seem not enhance viral transmission within the organism as JCPyV viremia is rarely observed in JCPVAN patients (3, 4, 217).

The role of viral mutations in the pathogenesis of BKPVAN seems similar to JCPVAN. Archetype BKPyV strains, thought to persist in the kidneys (10), represented the major viral population in the plasma and urine of all but one kidney transplant patient with or without clinical and histological evidence for BKPVAN. Contrast to the archetype-like strain of one PML patient, the archetype-like strain associated with BKPVAN had decreased number of binding sites for transcriptional-activator NF-1 suggesting either that JCPyV and BKPyV have different host cells in the kidneys or different transcription factors regulate their replication in a same cell type. Similar to the previous studies (10, 122, 206), also small populations of mutated BKPyV strains were detected in the urine and plasma of patients with or without BKPVAN; robust viral replication in the kidneys may allow spreading of viral particles via the systemic circulation. Further, mutated BKPyV strains, albeit in minority, might have altered pathogenic capacity due to enhancement of viral microRNA expression.

No curative or prophylactic treatment for PML or polyomavirus-associated nephropathy currently exists leaving the monitoring and early identification of patients at risk as the only option. At present, the only way to evaluate PML risk is to screen for the presence of JCPyV-specific antibodies. This way PML risk can however be excluded only in JCPyV seronegative patients representing approximately 50% of cases (153, 154). Moreover, among JCPyV seropositive patients only very few will finally develop PML. More accurate diagnostic methods for monitoring of JCPyV reactivation is thus needed. Although kidney transplant patients are routinely monitored for the presence of BKPyV replication (189), the prognosis of BKPVAN remains problematic. No monitoring of kidney transplant patients for JCPVAN is currently recommended due to rarity of the disease (218). Regardless of the small cohort size and retrospective nature of this thesis, it seems that monitoring for the emergence of neurotropic JCPyV strains could be used to predict the development of PML. On the contrary, the presence of mutated BKPyV strains seems not be associated with BKPVAN onset but they are rather a by-product of robust viral replication in the kidneys.

With advances in novel molecular techniques, such as MPS, our knowledge on human polyomaviruses is only expected to increase. Apart from prolonged immunosuppression, predisposing factors of human polyomavirus-associated diseases remain largely unresolved. MPS enables extremely detailed characterization of viral strains present in a biological sample providing a novel way for investigating the biology of human polyomaviruses that is important in order to improve the prognosis and prevention of polyomavirus-associated diseases.

5 CONCLUDING REMARKS

Human polyomaviruses are important human pathogens causing various clinical manifestations almost exclusively in immunosuppressed individuals. In the last decade, the number of known human polyomaviruses has expanded mostly owing to the establishment of highly sensitive MPS techniques allowing the characterization of very small viral populations present in a sample. Further, the PML risk associated with the use of some novel biological drugs has increased the interest in the biology of JCPyV, discovered already in the 1970s. This thesis summarizes the studies investigating the sequence variation of JCPyV and BKPyV in association with PML, BKPVAN, and JCPVAN: how mutations occurring within the viral genome could be associated with the disease pathogenesis (**Figure 14**).

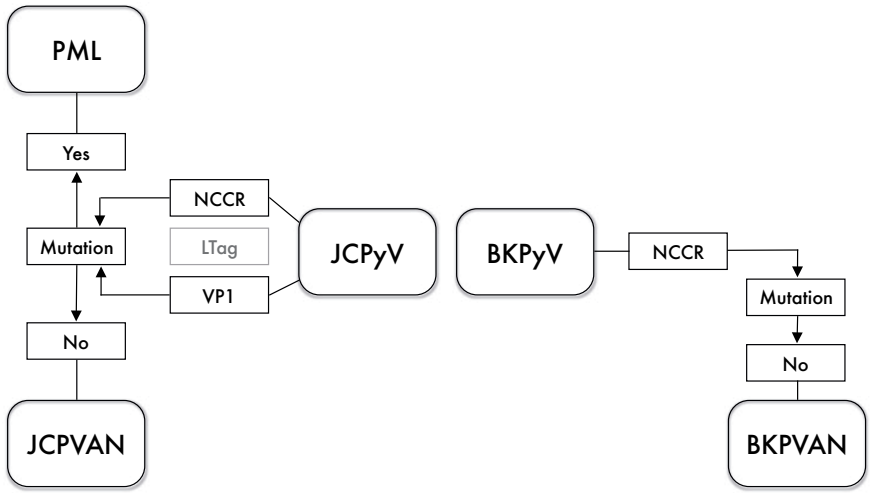


Figure 14. Sequence variation of JCPyV and BKPyV and its role in the pathogenesis of PML, JCPVAN, and BKPVAN.

In this work, complete JCPyV genomes of PML patients and kidney transplant patients with either stable graft function or histologically confirmed JCPVAN were characterized. In agreement with previous studies, neurotropic strains harboring rearrangements within the NCCR were shown to associate with PML; moreover, for the first time multiple neurotropic strains were shown to be associated with PML. The presence of archetype-like strain in one patient indicated, however, that NCCR rearrangements are not always necessary for the disease onset and could be compensated with mutations occurring in the protein coding region such as in the major capsid protein VP1. In contrast to PML, mutations within JCPyV genome seem not necessary for the development of JCPVAN since archetype or archetype-like JCPyV strains dominated in urine of all kidney transplant patients. Other factors, such as presence of JCPyV-specific antibodies, could play a role in the pathogenesis of this kidney disease.

To study the role of viral mutations in the development of BKPVAN, BKPyV TCR regions of kidney transplant patients with or without clinical or histological evidence of BKPVAN were characterized. Similar to JCPVAN, archetype BKPyV predominated in all but one patient with biopsy-confirmed BKPVAN. Small populations of rearranged strains were also identified in most patients suggesting that enhanced viral replication

in the kidneys due to immunosuppression could allow accumulation of mutations within the NCCR and emergence of rearranged strains that, even as a minority, could play a role in the viral pathogenesis by affecting e.g. the expression of BKPyV-encoded microRNAs.

PML and polyomavirus-associated nephropathy are severe, even fatal, chronic diseases. No prophylactic or curative treatment for these diseases currently exists leaving the monitoring of patients as the only option. This thesis has increased the knowledge on the role of sequence variation in the pathogenicity of JCPyV and BKPyV and development of PML, BKPVAN, and JCPVAN. The results provide new information to be used in the establishment of more sensitive and accurate diagnostic methods for the prognosis of these polyomavirus-associated diseases.

6 ACKNOWLEDGEMENTS

This thesis was a joint effort and would not have been finished without help from many. This thesis was carried out in the Department of Virology and Immunology, Helsinki University Hospital Laboratory, and Department of Virology, Faculty of Medicine, University of Helsinki. I want to acknowledge Director Maija Lappalainen (Division of Clinical Microbiology, Helsinki University Hospital Laboratory) and the head of the department, Professor Kalle Saksela (University of Helsinki) for providing excellent facilities for my research work.

First and foremost, I want to thank my supervisor, Docent Eeva Auvinen. I admire your endless enthusiasm for science, in particular for virology, that has served as an inspiration for this thesis. Thank you for your support, excellent advices, and for giving me the freedom to develop as an independent researcher. Thank you for believing in me in times when I myself have not. The former group member, PhD Elina Virtanen is warmly thanked for her support, both mental and science-related, during this thesis work. Thank you for all those long but fruitful brainstorming sessions. My thesis committee, Emeritus Professor Veijo Hukkanen, Professor Petri Auvinen, and Docent Sisko Tauriainen are also acknowledged for providing critical insights and support during this research work.

I want to warmly thank all our excellent collaborators who have been participating in this thesis work. I want to acknowledge all members of the DNA genomics and sequencing lab, at the Institute of Biotechnology, University of Helsinki for their excellent assistance. Special thanks go to MSc Pia Laine and Laboratory Engineer Lars Paulín for their endless support and guidance in the sequencing-related matters. This thesis would not have been finished without you. The members of PCR lab at the Department of Virology and Immunology, Helsinki University Hospital Laboratory are warmly thanked for their aid in countless practical issues. I want to also express my gratitude to my dear PMS lab members, Kati Hokynar, Anu Kaitonen, and Satu Hänninen for surrounding me with cheerful, and sometimes crazy, atmosphere.

Last but by far not the least, I want to thank my family for being always there for me. I am really fortunate to have you all in my life. I am also eternally grateful for my beloved husband, Niclas. Thank you for your endless patience and support during this project. I love you.

This thesis was funded by Juhani Ahon Lääketieteen Tutkimussäätiö, Munuaissäätiö, Suomen MS-säätiö, and Infektiotautien Tutkimusyhdistys.

Helsinki, November 2019

Hanna Liimatainen

7 APPENDICES

Supplementary Table 1. Characteristics of PacBio long-read sequence data obtained from three PML patients, twenty asymptomatic kidney transplant patients, and one kidney transplant patient with histologically confirmed JCPVAN.

Study	Patient (sample)	Subreads	Base coverage	Strain	Subread coverage	Base pair coverage	CCS length (bp)	Accuracy (%)
I	1 (CSF)	17,260	49,293,033	1.1	38	713.91	5,127	98.828
				1.2	159	1,142.55	5,233	99.955
				1.3	142	1,012.33	5,285	99.994
	2 (CSF)	32,537	99,154,011	2.1	>500	4,380.96	5,151	99.994
	3 (CSF)	25,537	83,007,590	3.1	>500	4,787.63	5,176	99.994
II	4 (urine)	16,200	49,100,679	4.1	>500	2,933.94	5,163	99.994
	5 (urine)	17,175	53,137,802	5.1	>500	3,214.79	5,163	99.994
	6 (urine)	22,634	63,755,490	6.1	>500	2,689.82	5,163	99.994
	7 (urine)	18,022	54,257,074	7.1	>500	3,201.02	5,163	99.994
	8 (urine)	14,770	45,942,614	8.1	>500	3,064.2	5,163	99.994
	9 (urine)	16,149	45,926,868	9.1	438	2,431.03	5,164	99.981
	10 (urine)	20,646	60,102,923	10.1	>500	3,261.35	5,164	99.994
	11 (urine)	26,747	90,483,697	11.1	>500	5,476.45	5,164	99.994
	12 (urine)	22,054	67,727,031	12.1	190	1,500.61	5,164	99.974
				12.2	256	2,191.28	5,151	99.994
	13 (urine)	26,577	82,484,521	13.1	>500	4,557.91	5,163	99.994
	14 (urine)	25,990	81,711,239	14.1	>500	3,924.82	5,149	99.994
	15 (urine)	22,876	71,859,179	15.1	>500	3,868.76	5,161	99.994
	16 (urine)	27,911	82,797,459	16.1	203	2,385.37	5,163	99.994
				16.2	297	1,402.63	5,164	99.984
	17 (urine)	16,497	51,942,003	17.1	>500	3,281.13	5,162	99.994
	18 (urine)	18,926	82,484,521	18.1	214	1,605.64	5,163	99.993
				18.2	286	2,248.46	5,162	99.994
	19 (urine)	14,077	43,397,292	19.1	432	2,815.16	5,163	99.994
	20 (urine)	20,646	60,102,923	20.1	>500	3,885.83	5,164	99.994
	21 (urine)	9,060	29,789,576	21.1	239	2,912.89	5,163	99.994
				21.2	261	2,932.18	5,163	99.994
	22 (urine)	10,108	29,958,288	22.1	253	2,323.65	5,162	99.994
				22.2	247	2,940.61	5,161	99.986
	23 (urine)	10,394	33,779,743	23.1	>500	6,122.65	5,162	99.994
	24 (plasma 1)	1,842	5,479,907	N/A	N/A	N/A	N/A	N/A
	24 (plasma 2)	3,529	10,325,670	N/A	N/A	N/A	N/A	N/A
	24 (plasma 3)	6,180	17,613,781	N/A	N/A	N/A	N/A	N/A
	24 (urine 1)	5,008	15,801,263	24.1	>500	3,326.04	5,164	99.994
	24 (urine 2)	6,346	20,574,684	24.2	>500	3,769.44	5,163	99.994
	24 (urine 3)	5,875	18,458,573	24.3	>500	3,502.83	5,164	99.994
	24 (urine 4)	22,212	79,665,903	24.4	>500	13,954.95	5,162	99.994

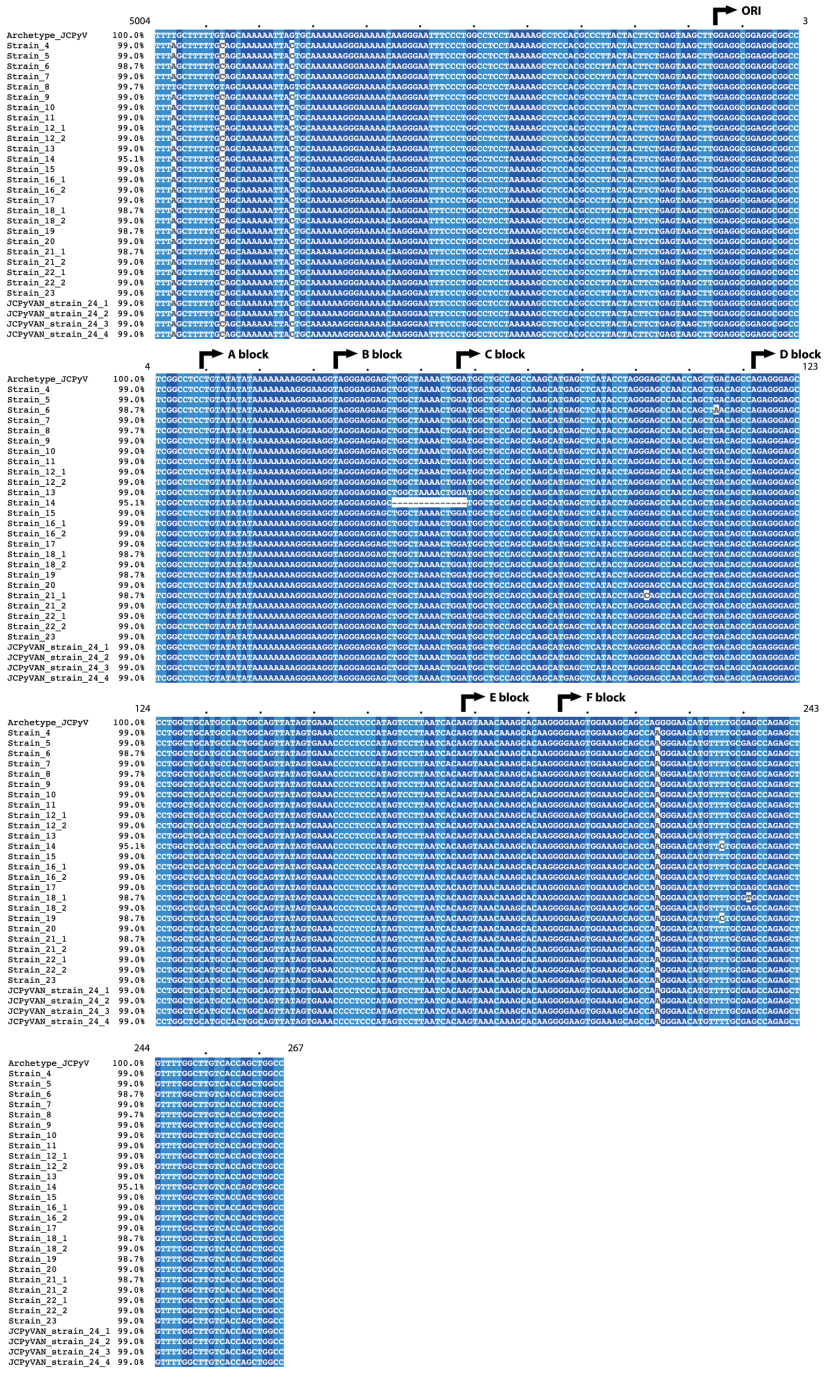
Supplementary Table 2. Genotypes and protein coding region mutations in the JCPyV strains of PML patients and kidney transplant patients with stable graft function or histologically confirmed JCPVAN. Nucleotide identity in the coding region between the patient strains and the reference genotype strain is shown.

Patient	Status	Strain	Genome (bp)	Genotype (subtype)	Nucleotide identity in the coding region (%)
1	PML	1.1	5,051	1 (A)	4836/4842 (99)
		1.2	5,191	1 (A)	4836/4841 (99)
		1.3	5,243	1 (A)	4836/4841 (99)
2	PML	2	5,111	1 (A)	4834/4854 (99)
3	PML	3	5,134	2 (B)	4845/4854 (99)
4	KTx	4	5,121	1 (A)	4837/4854 (99)
5	KTx	5	5,121	1 (A)	4837/4854 (99)
6	KTx	6	5,121	1 (A)	4835/4854 (99)
7	KTx	7	5,121	1 (B)	4839/4854 (99)
8	KTx	8	5,121	2 (B)	4851/4854 (99)
9	KTx	9	5,121	1 (A)	4834/4855 (99)
10	KTx	10	5,122	4	4842/4855 (99)
11	KTx	11	5,122	4	4841/4855 (99)
12	KTx	12.1	5,122	1 (B)	4841/4855 (99)
		12.2	5,109	1 (B)	4829/4854 (99)
13	KTx	13	5,121	1 (B)	4838/4854 (99)
14	KTx	14	5,107	1 (A)	4837/4854 (99)
15	KTx	15	5,121	1 (B)	4841/4854 (99)
16	KTx	16.1	5,121	1 (B)	4841/4854 (99)
		16.2	5,122	1 (B)	4839/4855 (99)
17	KTx	17	5,121	1 (B)	4840/4854 (99)
18	KTx	18.1	5,122	1 (A)	4850/4855 (99)
		18.2	5,121	1 (A)	4850/4854 (99)
19	KTx	19	5,121	1 (A)	4837/4854 (99)
20[b]	KTx	20	5,122	1 (B)	4838/4855 (99)
21	KTx	21.1	5,121	1 (B)	4841/4854 (99)
		21.2	5,121	1 (B)	4841/4854 (99)
22	KTx	22.1	5,121	1 (B)	4839/4854 (99)
		22.2	5,121	1 (B)	4838/4854 (99)
23	KTx	23	5,121	1 (B)	4840/4854 (99)
24	JCPVAN	24.1-24.4	5,122	4	4840/4855 (99)

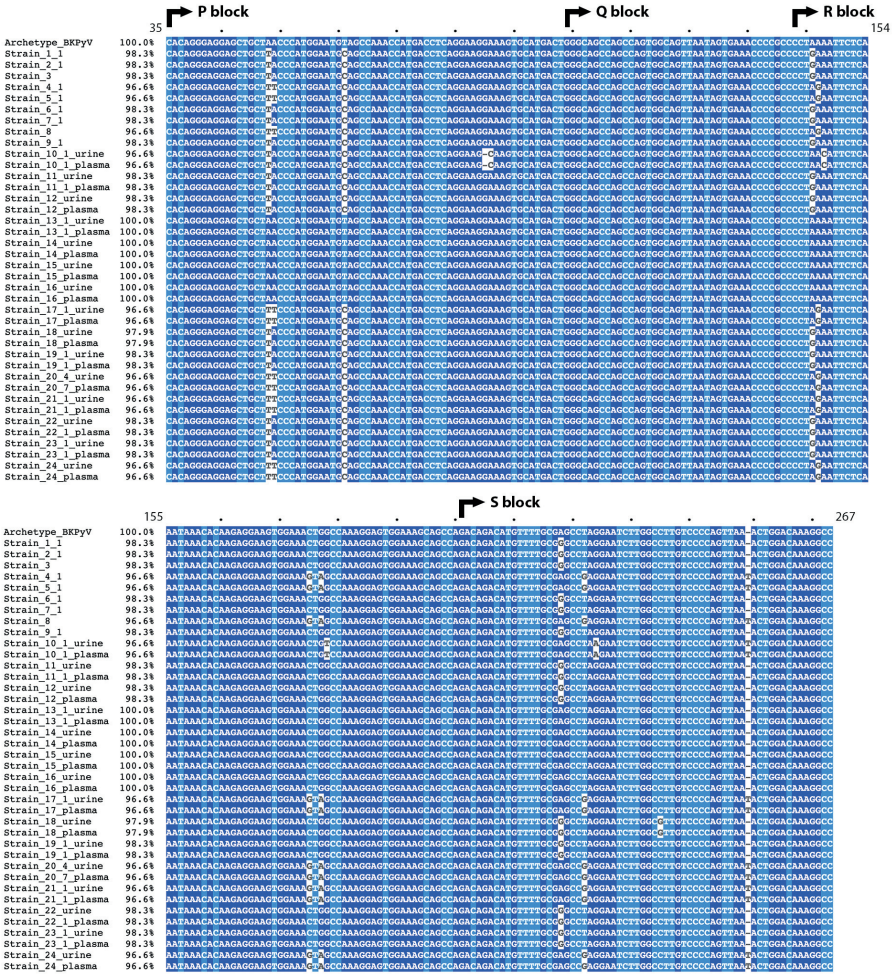
Abbreviations: PML, progressive multifocal leukoencephalopathy; KTx, kidney transplantation; JCPVAN, JC polyomavirus-associated nephropathy; bp, base pair.

Supplementary Table 3. Characteristics of the short-read sequence data from twenty-four kidney transplant patients with or without biopsy-confirmed BKPVAN.

Study	Patient	Source	BKPyV load (copies/mL)	Original read amount	Post-analytic read amount	Proportion (%)
III	1	Plasma	4.53E+04	174,074	149,901	86.1
	2	Plasma	1.90E+05	170,929	150,401	88.0
	3	Plasma	2.67E+04	143,719	126,216	87.8
	4	Plasma	8.91E+04	154,513	133,708	86.5
	5	Plasma	2.17E+05	191,528	163,993	85.6
	6	Plasma	8.5E+03	184,170	151,869	82.5
	7	Plasma	2.96E+04	193,089	171,162	88.6
	8	Plasma	1.16E+05	217,372	187,630	86.3
	9	Plasma	1.93E+04	145,310	125,964	86.7
IV	10	Urine	3.10E+07	194,648	125,244	64.3
		Plasma	5.40E+03	195,179	124,213	63.6
	11	Urine	3.10E+08	173,797	121,460	69.9
		Plasma	2.00E+03	206,449	126,478	61.3
	12	Urine	2.50E+08	179,658	129,518	72.1
		Plasma	2.70E+04	147,503	85,959	58.3
	13	Urine	3.70E+07	197,551	134,458	68.1
		Plasma	1.00E+04	179,504	102,914	57.3
	14	Urine	1.00E+10	209,807	128,873	61.4
		Plasma	1.10E+04	197,996	137,845	69.6
	15	Urine	1.00E+10	156,320	109,510	70.1
		Plasma	1.40E+05	174,402	111,574	64.0
	16	Urine	1.10E+08	186,056	121,698	65.4
		Plasma	2.30E+04	210,522	139,579	66.3
	17	Urine	2.80E+08	205,336	131,536	64.1
		Plasma	1.90E+04	170,259	107,301	63.0
	18	Urine	5.10E+08	203,181	147,722	72.7
		Plasma	3.80E+05	201,098	142,933	71.1
	19	Urine	1.00E+10	171,575	118,488	69.1
		Plasma	1.10E+05	197,232	135,637	68.8
	20	Urine	1.90E+11	289,372	217,378	75.1
		Plasma	4.20E+08	322,959	231,178	71.6
	21	Urine	1.60E+11	293,225	233,812	79.7
		Plasma	1.10E+07	352,815	262,639	74.4
	22	Urine	2.90E+10	294,844	237,925	80.7
		Plasma	4.70E+05	360,608	263,356	73.0
	23	Urine	2.00E+11	308,733	251,022	81.3
		Plasma	1.70E+06	322,959	231,178	71.6
	24	Urine	6.30E+07	288,023	230,811	80.1
		Plasma	1.40E+04	281,712	217,932	77.4



Supplementary Figure 1. Archetype and archetype-like JCPyV NCCR regions from kidney transplant patients with and without JCPVAN. JCPyV WW strain (AB038249) was used as an archetype reference strain. For multiple sequence alignment, the multiple sequence alignment tool Kalign and multiple sequence alignment viewer MView (EMBL-EBI) with default parameters were used.



Supplementary Figure 2. Archetype BKPyV TCR regions from kidney transplant patients with and without BKPVAN. BKPyV WW strain (AB211371.1) was used as an archetype reference strain. For multiple sequence alignment, the multiple sequence alignment tool Kalign and multiple sequence alignment viewer MView (EMBL-EBI) with default parameters were used.

8 REFERENCES

1. Padgett BL, Walker DL, Zurhein GM, Eckroade RJ, Dessel BH. **1971**. Cultivation of papova-like virus from human brain with progressive multifocal leucoencephalopathy. Lancet 1:1257-60.
2. Gardner SD, Field AM, Coleman DV, Hulme B. **1971**. New human papovavirus (BK) isolated from urine after renal transplantation. Lancet 1:1253-7.
3. Yang D, Keys B, Conti DJ, Foulke L, Stellrecht K, Cook L, Lopez-Soler RI. **2017**. JC polyomavirus nephropathy, a rare cause of transplant dysfunction: Case report and review of literature. Transpl Infect Dis 19:e12654.
4. Kantarci G, Eren Z, Demirag A, Dogan I, Cakalagaoglu F, Yilmaz G. **2011**. JC virus-associated nephropathy in a renal transplant recipient and comparative analysis of previous cases. Transpl Infect Dis 13:89-92.
5. Feng HC, Shuda M, Chang Y, Moore PS. **2008**. Clonal integration of a polyomavirus in human Merkel cell carcinoma. Science 319:1096-1100.
6. van der Meijden E, Janssens RWA, Lauber C, Bavinck JNB, Gorbalenya AE, Feltkamp MCW. **2010**. Discovery of a new human polyomavirus associated with Trichodysplasia spinulosa in an immunocompromized patient. Plos Pathogens 6:e1001024.
7. Sunyaev SR, Lugovskoy A, Simon K, Gorelik L. **2009**. Adaptive mutations in the JC virus protein capsid are associated with progressive multifocal leukoencephalopathy (PML). Plos Genetics 5:e1000268.
8. Gosert R, Kardas P, Major EO, Hirsch HH. **2010**. Rearranged JC virus noncoding control regions found in progressive multifocal leukoencephalopathy patient samples increase virus early gene expression and replication rate. J Virol 84:10448-10456.
9. Zheng HY, Ikegaya H, Takasaka T, Matsushima-Ohno T, Sakurai M, Kanazawa I, et al. **2005**. Characterization of the VP1 loop mutations widespread among JC polyomavirus isolates associated with progressive multifocal leukoencephalopathy. Biochem Biophys Res Commun 333:996-1002.
10. Gosert R, Rinaldo CH, Funk GA, Egli A, Ramos E, Drachenberg CB, Hirsch HH. **2008**. Polyomavirus BK with rearranged noncoding control region emerge in vivo in renal transplant patients and increase viral replication and cytopathology. J Exp Med 205:841-852.
11. Metzker ML. **2010**. Sequencing technologies: The next generation. Nature Reviews Genetics 11:31-46.
12. Zagordi O, Klein R, Daumer M, Beerenwinkel N. **2010**. Error correction of next-generation sequencing data and reliable estimation of HIV quasispecies. Nucleic Acids Res 38:7400-7409.
13. Fishman JA. **2007**. Medical progress: Infection in solid-organ transplant recipients. New Engl J Med 357:2601-2614.
14. French MA, Price P, Stone SF. **2004**. Immune restoration disease after antiretroviral therapy. AIDS 18:1615-1627.
15. Morrison VA, Dunn DL, Manivel JC, Gajlpeczalska KJ, Peterson BA. **1994**. Clinical characteristics of posttransplant lymphoproliferative disorders. Am J Med 97:14-24.
16. Broers AEC, van der Holt R, van Esser JWJ, Gratama JW, Henzen-Logmans S, Kuenen-Boumeester V, Lowenberg B, Cornelissen JJ. **2000**. Increased transplant-related morbidity and mortality in CMV-seropositive patients despite highly effective prevention of CMV disease after allogeneic T-cell-depleted stem cell transplantation. Blood 95:2240-2245.
17. Beral V, Peterman TA, Berkelman RL, Jaffe HW. **1990**. Kaposi's sarcoma among

persons with AIDS - a sexually-transmitted infection. Lancet 335:123-128.
18. Moens U, Calvignac-Spencer S, Lauber C, Ramgvist T, Feltkamp MCW, Daugherty MD, Verschoor EJ, Ehlers B, Consortium IR. **2017**. ICTV Virus Taxonomy Profile: Polyomaviridae. J Gen Virol 98:1159-1160.
19. Dang-Tan T, Mahmud SM, Puntoni R, Franco EL. **2004**. Polio vaccines, Simian Virus 40, and human cancer: The epidemiologic evidence for a causal association. Oncogene 23:6535-6540.
20. Polyomaviridae Study Group of the International Committee on Taxonomy of Viruses, Calvignac-Spencer S, Feltkamp MC, Daugherty MD, Moens U, Ramqvist T, Johne R, Ehlers B. **2016**. A taxonomy update for the family Polyomaviridae. Arch Virol 161:1739-1750.
21. Gheit T, Dutta S, Oliver J, Robitaille A, Hampras S, Combes JD, McKay-Chopin S, Le Calvez-Kelm F, et al. **2017**. Isolation and characterization of a novel putative human polyomavirus. Virology 506:45-54.
22. Kamminga S, van der Meijden E, Feltkamp MCW, Zaaijer HL. **2018**. Seroprevalence of fourteen human polyomaviruses determined in blood donors. PloS One 12:e0206273.
23. Kean JM, Rao S, Wang M, Garcea RL. **2009**. Seroepidemiology of human polyomaviruses. Plos Pathogens 5:e1000363.
24. Gossai A, Waterboer T, Nelson HH, Michel A, Willhauck-Fleckenstein M, Farzan SF, Hoen AG, Christensen BC, et al. **2016**. Seroepidemiology of human polyomaviruses in a US population. Am J Epidemiol 183:61-69.
25. Nicol JTJ, Leblond V, Arnold F, Guerra G, Mazzoni E, Tognon M, Coursaget P, Touze A. **2014**. Seroprevalence of human Malawi polyomavirus. J Clin Microbiol 52:321-323.
26. Gaboriaud P, Ferte M, Arnold F, Leblond V, Nicol J, Debare H, Le Meur M, Martini F, et al. **2018**. Age-specific seroprevalence of human polyomavirus 12 and Saint Louis and New Jersey polyomaviruses. Emerg Microbes Infect 7:22.
27. Nicol JTJ, Robinot R, Carpentier A, Carandina G, Mazzoni E, Tognon M, Touze A, Coursaget P. **2013**. Age-specific seroprevalences of Merkel cell polyomavirus, human polyomaviruses 6, 7, and 9, and Trichodysplasia spinulosa-associated polyomavirus. Clin Vaccine Immunol 20:363-368.
28. Lim ES, Meinerz NM, Primi B, Wang D, Garcea RL. **2014**. Common exposure to STL polyomavirus during childhood. Emerg Infect Dis 20:1559-1561.
29. Moens U, Van Ghelue M, Song XB, Ehlers B. **2013**. Serological cross-reactivity between human polyomaviruses. Rev Med Virol 23:250-264.
30. Comar M, Zanotta N, Bovenzi M, Campello C. **2010**. JCV/BKV and SV40 viral load in lymphoid tissues of young immunocompetent children from an area of North-East Italy. J Med Virol 82:1236-1240.
31. Monaco MCG, Atwood WJ, Gravell M, Tornatore CS, Major EO. **1996**. JC virus infection of hematopoietic progenitor cells, primary B lymphocytes, and tonsillar stromal cells: Implications for viral latency. J Virol 70:7004-7012.
32. Bialasiewicz S, Whiley DM, Lambert SB, Nissen MD, Sloots TP. **2009**. Detection of BK, JC, WU, or KI polyomaviruses in faecal, urine, blood, cerebrospinal fluid and respiratory samples. J Clin Virol 45:249-254.
33. Allander T, Andreasson K, Gupta S, Bjerkner A, Bogdanovic G, Persson MAA, Dalianis T, Ramqvist T, et al. **2007**. Identification of a third human polyomavirus. J Virol 81:4130-4136.
34. Gaynor AM, Nissen MD, Whiley DM, Mackay IM, Lambert SB, Wu G, Brennan DC, Storch GA, et al. **2007**. Identification of a novel polyomavirus from patients with acute respiratory tract infections. Plos Pathogens 3:595-604.
35. Rockett RJ, Sloots TP, Bowes S, O'Neill N, Ye SF, Robson J, Whiley DM, Lambert SB,

et al. **2013**. Detection of novel polyomaviruses, TSPyV, HPyV6, HPyV7, HPyV9 and MWPyV in feces, urine, blood, respiratory swabs and cerebrospinal fluid. Plos One 8:e62764.
36. Korup S, Rietscher J, Calvignac-Spencer S, Trusch F, Hofmann J, Moens U, Sauer I, Voigt S, et al. **2013**. Identification of a novel human polyomavirus in organs of the gastrointestinal tract. Plos One 8:58021.
37. Lim ES, Reyes A, Antonio M, Saha D, Ikumapayi UN, Adeyemi M, Stine OC, Skelton R, et al. **2013**. Discovery of STL polyomavirus, a polyomavirus of ancestral recombinant origin that encodes a unique T antigen by alternative splicing. Virology 436:295-303.
38. Foulongne V, Sauvage V, Hebert C, Dereure O, Cheval J, Gouilh MA, Pariente K, Segondy M, et al. **2012**. Human skin microbiota: High diversity of DNA viruses identified on the human skin by high throughput sequencing. Plos One 7:38499.
39. Mishra N, Pereira M, Rhodes RH, An P, Pipas JM, Jain K, Kapoor A, Briese T, et al. **2014**. Identification of a novel polyomavirus in a pancreatic transplant recipient with retinal blindness and vasculitic myopathy. J Infect Dis 210:1595-1599.
40. Griffith JP, Griffith DL, Rayment I, Murakami WT, Caspar DLD. **1992**. Inside polyomavirus at 25-Å resolution. Nature 355:652-654.
41. Calvignac-Spencer S, Feltkamp MCW, Daugherty MD, Moens U, Ramqvist T, Johne R, Ehlers B, Int Comm Taxonomy V. **2016**. A taxonomy update for the family Polyomaviridae. Arch Virol 161:1739-1750.
42. Carter JJ, Daugherty MD, Qi XJ, Bheda-Malge A, Wipf GC, Robinson K, Roman A, Malik HS, Galloway DA. **2013**. Identification of an overprinting gene in Merkel cell polyomavirus provides evolutionary insight into the birth of viral genes. Proc Natl Acad Sci U S A 110:12744-12749.
43. van der Meijden E, Kazem S, Dargel CA, van Vuren N, Hensbergen PJ, Feltkamp MCW. **2015**. Characterization of T antigens, Including middle T and alternative T, expressed by the human polyomavirus associated with Trichodysplasia spinulosa. J Virol 89:9427-9439.
44. Bialasiewicz S, Whiley DM, Lambert SB, Jacob K, Bletchly C, Wang D, Nissen MD, Sloots TP. **2008**. Presence of the newly discovered human polyomaviruses KI and WU in Australian patients with acute respiratory tract infection. J Clin Virol 41:63-68.
45. Cheng JW, Rozenblatt-Rosen O, Paulson KG, Nghiem P, DeCaprio JA. **2013**. Merkel cell polyomavirus large T antigen has growth-promoting and inhibitory activities. J Virol 87:6118-6126.
46. Schowalter RM, Buck CB. **2013**. The Merkel cell polyomavirus minor capsid protein. Plos Pathogens 9:e1003558.
47. Neu U, Maginnis MS, Palma AS, Stroh LJ, Nelson CDS, Feizi T, Atwood WJ, Stehle T. **2010**. Structure-function analysis of the human JC polyomavirus establishes the LSCT pentasaccharide as a functional receptor motif. Cell Host Microbe 8:309-319.
48. Neu U, Allen SAA, Blaum BS, Liu Y, Frank M, Palma AS, Stroh LJ, Feizi T, et al. **2013**. A Structure-guided mutation in the major capsid protein retargets BK polyomavirus. Plos Pathogens 9:e1003688.
49. Gee GV, Tsomaia N, Mierke DF, Atwood WJ. **2004**. Modeling a sialic acid binding pocket in the external loops of JC virus VP1. J Biol Chem 279:49172-49176.
50. Ajuh ET, Wu ZS, Kraus E, Weissbach FH, Bethge T, Gosert R, Fischer N, Hirsch HH. **2018**. Novel human polyomavirus noncoding control regions differ in bidirectional gene expression according to host cell, large T-antigen expression, and clinically occurring rearrangements. J Virol 92:e02231-17.
51. Wu JH, Nguyen HP, Rady PL, Tyring SK. **2016**. Molecular insight into the viral

biology and clinical features of Trichodysplasia spinulosa. Br J Dermatol 174:490-498.
52. van der Meijden E, Horvath B, Nijland M, de Vries K, Raez E, Diercks GF, de Weerd AE, Clahsen-van Groningen MC, et al. **2017**. Primary polyomavirus infection, not reactivation, as the cause of Trichodysplasia spinulosa in immunocompromised patients. J Infect Dis 215:1080-1084.
53. van Veenendaal LM, van Akkooi ACJ, Verhoef C, Grunhagen DJ, Klop WMC, Valk GD, Tesselaar MET. **2018**. Merkel cell carcinoma: Clinical outcome and prognostic factors in 351 patients. J Surg Oncol 117:1768-1775.
54. Fitzgerald TL, Dennis S, Kachare SD, Vohra NA, Wong JH, Zervos EE. **2015**. Dramatic increase in the incidence and mortality from Merkel cell carcinoma in the United States. Am Surg 81:802-806.
55. Moshiri AS, Doumani R, Yelistratova L, Blom A, Lachance K, Shinohara MM, Delaney M, Chang O, et al. **2017**. Polyomavirus-negative Merkel cell carcinoma: A more aggressive subtype based on analysis of 282 cases using multimodal tumor virus detection. J Invest Dermatol 137:819-827.
56. Houben R, Shuda M, Weinkam R, Schrama D, Feng HC, Chang YA, Moore PS, Becker JC. **2010**. Merkel cell polyomavirus-infected Merkel cell carcinoma cells require expression of viral T antigens. J Virol 84:7064-7072.
57. Shuda M, Feng HC, Kwun HJ, Rosen ST, Gjoerup O, Moore PS, Chang Y. **2008**. T antigen mutations are a human tumor-specific signature for Merkel cell polyomavirus. Proc Natl Acad Sci U S A 105:16272-16277.
58. Nguyen KD, Lee EE, Yue Y, Stork J, Pock L, North JP, Vandergriff T, Cockerell C, et al. **2016**. Human polyomavirus 6 and 7 Are associated with a pruritic and dyskeratotic dermatosis. J Am Acad Dermatol 76:932-940.
59. Payungporn S, Chieochansin T, Thongmee C, Samransammajkit R, Theamboolers A, Poovorawan Y. **2008**. Prevalence and molecular characterization of WU/KI polyomaviruses isolated from pediatric patients with respiratory disease in Thailand. Virus Res 135:230-236.
60. Ma FL, Li DD, Wei TL, Li JS, Zheng LS. **2017**. Quantitative detection of human Malawi polyomavirus in nasopharyngeal aspirates, sera, and feces in Beijing, China, using real-time TaqMan-based PCR. Virology J 14:152.
61. Nguyen KD, Lee EE, Yue YB, Stork J, Pock L, North JP, Vandergriff T, Cockerell C, et al. **2017**. Human polyomavirus 6 and 7 are associated with pruritic and dyskeratotic dermatoses. J Am Acad Dermatol 76:932-40.
62. Kassem A, Schopflin A, Diaz C, Weyers WG, Stickeler E, Werner M, zur Hausen A. **2008**. Frequent detection of Merkel cell polyomavirus in human Merkel cell carcinomas and identification of a unique deletion in the VP1 gene. Cancer Res 68:5009-5013.
63. Martin JD, King DM, Slauch JM, Frisque RJ. **1985**. Differences in regulatory sequences of naturally-occurring JC-virus variants. J Virol 53:306-311.
64. Grinnell BW, Padgett BL, Walker DL. **1983**. Comparison of infectious JC virus DNAs cloned from human-brain. J Virol 45:299-308.
65. Yogo Y, Kitamura T, Sugimoto C, Ueki T, Aso Y, Hara K, Taguchi F. **1990**. Isolation of a possible archetypal JC virus DNA sequence from nonimmunocompromised individuals. J Virol 64:3139-3143.
66. Rosen S, Harmon W, Krensky AM, Edelson PJ, Padgett BL, Grinnell BW, Rubino MJ, Walker DL. **1983**. Tubulo-interstitial nephritis associated with polyomavirus (BK type) infection. N Engl J Med 308:1192-1196.
67. Moens U, Johansen T, Johnsen JI, Seternes OM, Traavik T. **1995**. Noncoding control region of naturally-occurring BK virus variants - sequence comparison and functional-analysis. Virus Genes 10:261-275.

68. Fang CY, Shen CH, Wang ML, Chen PL, Chan MWY, Hsu PH, Chang DC. **2015**. Global profiling of histone modifications in the polyomavirus BK virion minichromosome. Virology 483:1-12.
69. Raj GV, Khalili K. **1995**. Transcriptional regulation: Lessons from the human neurotropic polyomavirus, JCV. Virology 213:283-291.
70. Sariyer IK, Khalili K. **2011**. Regulation of human neurotropic JC virus replication by alternative splicing factor SF2/ASF in glial cells. Plos One 6:e14630.
71. Liang B, Tikhanovich I, Nasheuer HP, Folk WR. **2012**. Stimulation of BK Virus DNA replication by NFI Family transcription factors. J Virol 86:3264-3275.
72. Akan I, Sariyer IK, Biffi R, Palermo V, Woolridge S, White MK, Amini S, Khalili K, et al. **2006**. Human polyomavirus JCV late leader peptide region contains important regulatory elements. Virology 349:66-78.
73. Cinque P, Koralnik IJ, Gerevini S, Miro JM, Price RW. **2009**. Progressive multifocal leukoencephalopathy in HIV-1 infection. Lancet Infect Dis 9:625-636.
74. An P, Robles MTS, Pipas JM. **2012**. Large T antigens of polyomaviruses: Amazing molecular machines. Annu Rev Microbiol 66:213-236.
75. Kelley WL, Georgopoulos C. **1997**. The T/t common exon of simian virus 40, JC, and BK polyomavirus T antigens can functionally replace the J-domain of the Escherichia coli DnaJ molecular chaperone. Proc Natl Acad Sci U S A 94:3679-3684.
76. Harris KF, Christensen JB, Imperiale MJ. **1996**. BK virus large T antigen: Interactions with the retinoblastoma family of tumor suppressor proteins and on cellular growth control. J Virol 70:2378-2386.
77. Bollag B, Prins C, Snyder EL, Frisque RJ. **2000**. Purified JC virus T and T ' proteins differentially interact with the retinoblastoma family of tumor suppressor proteins. Virology 274:165-178.
78. Meinke G, Phelan PJ, Kalekar R, Shin J, Archambault J, Bohm A, Bullock PA. **2014**. Insights into the initiation of JC virus DNA replication derived from the crystal structure of the T-antigen origin binding domain. Plos Pathogens 10:e1003966.
79. Deyerle KL, Sajjadi FG, Subramani S. **1989**. Analysis of origin of DNA-replication of human papovavirus BK. J Virol 63:356-365.
80. Gallia GL, Safak M, Khalili K. **1998**. Interaction of the single-stranded DNA-binding protein Purα with the human polyomavirus JC virus early protein T-antigen. The Journal of Biological Chemistry 273:32662-9
81. Safak M, Gallia GL, Ansari SA, Khalili K. **1999**. Physical and functional interaction between the Y-box binding protein YB-1 and human polyomavirus JC virus large T antigen. J Virol 73:10146-10157.
82. Bethge T, Hachemi HA, Manzetti J, Gosert R, Schaffner W, Hirsch HH. **2015**. Sp1 sites in the noncoding control region of BK polyomavirus are key regulators of bidirectional viral early and late gene expression. J Virol 89:3396-3411.
83. Pipas JM. **1992**. Common and unique features of T-antigens encoded by the polyomavirus group. J Virol 66:3979-3985.
84. Bollag B, Hofstetter CA, Reviriego-Mendoza MM, Frisque RJ. **2010**. JC virus small T antigen binds phosphatase PP2A and Rb family proteins and is required for efficient viral DNA replication activity. Plos One 5:e10606.
85. Pallas DC, Shahrik LK, Martin BL, Jaspers S, Miller TB, Brautigan DL, Roberts TM. **1990**. Polyoma small and middle T-antigens and SV40 small T-antigen form stable complexes with protein phosphatase-2A. Cell 60:167-176.
86. DeCaprio JA, Garcea RL. **2013**. A cornucopia of human polyomaviruses. Nature Rev Microbiol 11:264-276.
87. Trowbridge PW, Frisque RJ. **1995**. Identification of 3 new JC virus proteins generated by alternative splicing of the early viral messenger-RNA. J Neurovirol

1:195-206.
88. Abend JR, Joseph AE, Das D, Campbell-Cecen DB, Imperiale MJ. **2009**. A truncated T antigen expressed from an alternatively spliced BK virus early mRNA. J Gen Virol 90:1238-1245.
89. Safak M, Barrucco R, Darbinyan A, Okada Y, Nagashima K, Khalili K. **2001**. Interaction of JC virus Agno protein with T antigen modulates transcription and replication of the viral genome in glial cells. J Virol 75:1476-1486.
90. Gerits N, Johannessen M, Tummler C, Walquist M, Kostenko S, Snapkov I, van Loon B, Ferrari E, et al. **2015**. Agnoprotein of polyomavirus BK interacts with proliferating cell nuclear antigen and inhibits DNA replication. Virology J 12:7.
91. Suzuki T, Orba Y, Okada Y, Sunden Y, Kimura T, Tanaka S, Nagashima K, Hall WW, et al. **2010**. The human polyoma JC virus agnoprotein acts as a viroporin. Plos Pathog 6:1000801.
92. Panou MM, Prescott EL, Hurdiss DL, Swinscoe G, Hollinshead M, Caller LG, Morgan EL, Carlisle L, et al. **2018**. Agnoprotein is an essential egress factor during BK polyomavirus infection. Int J Mol Sci 19:902.
93. Li TC, Takeda N, Kato K, Nilsson J, Xing L, Haag L, Cheng RH, Miyamura T. **2003**. Characterization of self-assembled virus-like particles of human polyomavirus BK generated by recombinant baculoviruses. Virology 311:115-124.
94. Chang DC, Fung CY, Ou WC, Chao PC, Li SY, Wang ML, Huang YL, Tzeng TY, et al. **1997**. Self-assembly of the JC virus major capsid protein, VP1, expressed in insect cells. J Gen Virol 78:1435-1439.
95. Stehle T, Gamblin SJ, Yan YW, Harrison SC. **1996**. The structure of simian virus 40 refined at 3.1 Å resolution. Structure 4:165-182.
96. Neu U, Woellner K, Gauglitz G, Stehle T. **2008**. Structural basis of GM1 ganglioside recognition by simian virus 40. Proc Natl Acad Sci U S A 105:5219-5224.
97. Saribas AS, Mun S, Johnson J, El-Hajmoussa M, White MK, Safak M. **2014**. Human polyoma JC virus minor capsid proteins, VP2 and VP3, enhance large T antigen binding to the origin of viral DNA replication: Evidence for their involvement in regulation of the viral DNA replication. Virology 449:1-16.
98. Gasparovic ML, Gee GV, Atwood WJ. **2006**. JC virus minor capsid proteins VP2 and VP3 are essential for virus propagation. J Virol 80:10858-10861.
99. Bennett SM, Zhao LB, Bosard C, Imperiale MJ. **2015**. Role of a nuclear localization signal on the minor capsid Proteins VP2 and VP3 in BKPyV nuclear entry. Virology 474:110-116.
100. Seo GJ, Fink LHL, O'Hara B, Atwood WJ, Sullivan CS. **2008**. Evolutionarily conserved function of a viral microRNA. J Virol 82:9823-9828.
101. Bauman Y, Nachmani D, Vitenshtein A, Tsukerman P, Drayman N, Stern-Ginossar N, Lankry D, Gruda R, et al. **2011**. An identical miRNA of the human JC and BK polyomaviruses targets the stress-induced ligand ULBP3 to escape immune elimination. Cell Host Microbe 9:93-102.
102. Seo GJ, Chen CJ, Sullivan CS. **2009**. Merkel cell polyomavirus encodes a microRNA with the ability to autoregulate viral gene expression. Virology 383:183-187.
103. Broekema NM, Imperiale MJ. **2013**. miRNA regulation of BK polyomavirus replication during early infection. Proc Natl Acad Sci U S A 110:8200-8205.
104. Lagatie O, Van Loy T, Tritsmans L, Stuyver LJ. **2014**. Viral miRNAs in plasma and urine divulge JC polyomavirus infection. Virology J 11:158.
105. Tian YC, Li YJ, Chen HC, Wu HH, Weng CH, Chen YC, Lee CC, Chang MY, et al. **2014**. Polyomavirus BK-encoded microRNA suppresses autoregulation of viral replication. Biochem Biophys Res Commun 447:543-549.
106. Martelli F, Giannecchini S. **2017**. Polyomavirus microRNAs circulating in

biological fluids during viral persistence. Rev Med Virol 27:e1927.
107. Giovannelli I, Martelli F, Repice A, Massacesi L, Azzi A, Giannecchini S. **2015**. Detection of JCPyV microRNA in blood and urine samples of multiple sclerosis patients under natalizumab therapy. J Neurovirol 21:666-670.
108. Kim MH, Lee YH, Seo JW, Moon H, Kim JS, Kim YG, Jeong KH, Moon JY, et al. **2017**. Urinary exosomal viral microRNA as a marker of BK virus nephropathy in kidney transplant recipients. Plos One 12:20190068.
109. Li JYZ, McNicholas K, Yong TY, Rao N, Coates PTH, Higgins GD, Carroll RP, Woodman RJ, et al. **2014**. BK virus encoded microRNAs are present in blood of renal transplant recipients with BK viral nephropathy. Am J Transplant 14:1183-1190.
110. Jobes DV, Chima SC, Ryschkewitsch CF, Stoner GL. **1998**. Phylogenetic analysis of 22 complete genomes of the human polyomavirus JC virus. J Gen Virol 79:2491-2498.
111. Agostini HT, Ryschkewitsch CF, Brubaker GR, Shao J, Stoner GL. **1997**. Five complete genomes of JC virus Type 3 from Africans and African Americans. Arch Virol 142:637-655.
112. Agostini HT, Ryschkewitsch CF, Stoner GL. **1996**. Genotype profile of human polyomavirus JC excreted in urine of immunocompetent individuals. J Clin Microbiol 34:159-164.
113. Agostini HT, Ryschkewitsch CF, Mory R, Singer EJ, Stoner GL. **1997**. JC virus (JCV) genotypes in brain tissue from patients with progressive multifocal leukoencephalopathy (PML) and in urine from controls without PML: Increased frequency of JCV type 2 in PML. J Infect Dis 176:1-8.
114. Sugimoto C, Kitamura T, Guo J, AlAhdal MN, Shchelkunov SN, Otova B, Ondrejka P, Chollet JY, et al. **1997**. Typing of urinary JC virus DNA offers a novel means of tracing human migrations. Proc Natl Acad Sci U S A 94:9191-9196.
115. Jobes DV, Friedlaender JS, Mgone CS, Agostini HT, Koki G, Yanagihara R, Ng TCN, Chima SC, et al. **2001**. New JC virus (JCV) genotypes from Papua New Guinea and Micronesia (Type 8 and Type 2E) and evolutionary analysis of 32 complete JCV genomes. Arch Virol 146:2097-2113.
116. Jin L, Gibson PE, Knowles WA, Clewley JP. **1993**. BK virus antigenic variants : Sequence-analysis within the capsid VP1 epitope. J Med Virol 39:50-56.
117. Nishimoto Y, Zheng HY, Zhong S, Ikegaya H, Chen Q, Sugimoto C, Kitamura T, Yogo Y. **2007**. An Asian origin for subtype IV BK virus based on phylogenetic analysis. J Mol Evol 65:103-111.
118. Baksh FK, Finkelstein SD, Swalsky PA, Stoner GL, Ryschkewitsch CF, Randhawa P. **2001**. Molecular genotyping of BK and JC viruses in human polyomavirus-associated interstitial nephritis after renal transplantation. Am J Kidney Dis 38:354-365.
119. Zheng HY, Nishimoto Y, Chen Q, Hasegawa M, Zhong S, Ikegaya H, Ohno N, Sugimoto C, et al. **2007**. Relationships between BK virus lineages and human populations. Microbes Infect 9:204-213.
120. Varella RB, Zalona ACJ, Diaz NC, Zalis MG, Santoro-Lopes G. **2018**. BK polyomavirus genotypes Ia and Ib1 exhibit different biological properties in renal transplant recipients. Virus Res 243:65-68.
121. Monao MCG, Jensen PN, Hou J, Durham LC, Major EO. **1998**. Detection of JC virus DNA in human tonsil tissue: Evidence for site of initial viral infection. J Virol 72:9918-9923.
122. Dorries K, Vogel E, Gunther S, Czub S. **1994**. Infection of human polyomaviruses JC and BK in peripheral-blood leukocytes from immunocompetent individuals. Virology 198:59-70.

123. Komagome R, Sawa H, Suzuki T, Suzuki Y, Tanaka S, Atwood WJ, Nagashima K. **2002**. Oligosaccharides as receptors for JC virus. J Virol 76:12992-13000.
124. Elphick GF, Querbes W, Jordan JA, Gee GV, Eash S, Manley K, Dugan A, Stanifer M, et al. **2004**. The human polyomavirus, JCV, uses serotonin receptors to infect cells. Science 306:1380-1383.
125. Low JA, Magnuson B, Tsai B, Imperiale MJ. **2006**. Identification of gangliosides GD1b and GT1b as receptors for BK virus. J Virol 80:1361-1366.
126. Dugan AS, Eash S, Atwood WJ. **2005**. An N-linked glycoprotein with alpha-(2,3)-linked sialic acid is a receptor for BK virus. J Virol 79:14442-14445.
127. Pho MT, Ashok A, Atwood WJ. **2000**. JC virus enters human glial cells by clathrin-dependent receptor-mediated endocytosis. J Virol 74:2288-2292.
128. Eash S, Querbes W, Atwood WJ. **2004**. Infection of Vero cells by BK virus is dependent on Caveolae. J Virol 78:11583-11590.
129. Drachenberg CB, Papadimitriou JC, Wali R, Cubitt CL, Ramos E. **2003**. BK polyoma virus allograft nephropathy: Ultrastructural features from viral cell entry to lysis. Am J Transplant 3:1383-1392.
130. Wollebo HS, Woldemichaele B, Khalili K, Safak M, White MK. **2013**. Epigenetic regulation of polyomavirus JC. Virology J 10:264.
131. Chang CF, Wang ML, Fang CY, Chen PL, Wu SF, Chan MWY, Chang DC. **2011**. Analysis of DNA methylation in human BK virus. Virus Genes 43:201-207.
132. Chatterjee M, Weyandt TB, Frisque RJ. **2000**. Identification of archetype and rearranged forms of BK virus in leukocytes from healthy individuals. J Med Virol 60:353-362.
133. Elsner C, Dorries K. **1992**. Evidence of human polyomavirus BK and polyomavirus JC infection in normal brain-tissue. Virology 191:72-80.
134. Ambalathingal GR, Francis RS, Smyth MJ, Smith C, Khanna R. **2017**. BK polyomavirus: Clinical aspects, immune regulation, and emerging therapies. Clin Microbiol Rev 30:503-528.
135. Marzocchetti A, Tompkins T, Clifford DB, Gandhi RT, Kesari S, Berger JR, Simpson DM, Prosperi M, et al. **2009**. Determinants of survival in progressive multifocal leukoencephalopathy. Neurology 73:1551-1558.
136. Yousry TA, Pelletier D, Cadavid D, Gass A, Richert ND, Radue EW, Filippi M. **2012**. Magnetic resonance imaging pattern in natalizumab-associated progressive multifocal leukoencephalopathy. Ann Neurol 72:779-787.
137. Rhein GMZ, Chou SM. **1965**. Particles resembling papova viruses in human cerebral demyelinating disease. Science 148:1477-9.
138. Clifford DB, DeLuca A, Simpson DM, Arendt G, Giovannoni G, Nath A. **2010**. Natalizumab-associated progressive multifocal leukoencephalopathy in patients with multiple sclerosis: Lessons from 28 cases. Lancet Neurol 9:438-446.
139. Engsig FN, Hansen ABE, Omland LH, Kronborg G, Gerstoft J, Laursen AL, Pedersen C, Mogensen CB, et al. **2009**. Incidence, clinical presentation, and outcome of progressive multifocal leukoencephalopathy in HIV-infected patients during the highly active antiretroviral therapy era: A nationwide cohort study. J Infect Dis 199:77-83.
140. Mateen FJ, Muralidharan R, Carone M, van de Beek D, Harrison DM, Aksamit AJ, Gould MS, Clifford DB, et al. **2011**. Progressive multifocal leukoencephalopathy in transplant recipients. Ann Neurol 70:305-322.
141. Berger JR, Aksamit AJ, Clifford DB, Davis L, Koralnik IJ, Sejvar JJ, Bartt R, Major EO, Nath A. **2013**. PML diagnostic criteria consensus statement from the AAN Neuroinfectious Disease Section. Neurology 80:1430-1438.
142. Gheuens S, Pierone G, Peeters P, Koralnik IJ. **2010**. Progressive multifocal

leukoencephalopathy in individuals with minimal or occult immunosuppression. J Neurol Neurosurg Psychiatry 81:247-254.
143. Molloy ES, Calabrese LH. **2009**. Progressive multifocal leukoencephalopathy: A national estimate of frequency in systemic lupus erythematosus and other rheumatic diseases. Arthritis Rheum 60:3761-3765.
144. Kartau M, Verkkoniemi-Ahola A, Paetau A, Palomäki M, Janes R, Ristola M, Lappalainen M, Anttila VJ. **2019**. The incidence and predisposing factors of John Cunningham virus-induced progressive multifocal leukoencephalopathy in Southern Finland: A population-based study. Open Forum Infectious Diseases 6:ofz024.
145. Tada H, Rappaport J, Lashgari M, Amini S, Wongstaal F, Khalili K. **1990**. Transactivation of the JC virus late promoter by the tat protein of type-1 human-immunodeficiency-virus in glial-cells. Proc Natl Acad Sci U S A 87:3479-3483.
146. Garcia-Suarez J, de Miguel D, Krsnik I, Banas H, Arribas I, Burgaleta C. **2005**. Changes in the natural history of progressive multifocal leukoencephalopathy in HIV-negative lymphoproliferative disorders: Impact of novel therapies. Am J Hematol 80:271-281.
147. Amend KL, Turnbull B, Foskett N, Napalkov P, Kurth T, Seeger J. **2010**. Incidence of progressive multifocal leukoencephalopathy in patients without HIV. Neurology 75:1326-1332.
148. Calabrese LH, Molloy ES, Huang DR, Ransohoff RM. **2007**. Progressive multifocal leukoencephalopathy in rheumatic diseases - evolving clinical and pathologic patterns of disease. Arthritis Rheum 56:2116-2128.
149. Krumbholz M, Meinl I, Kumpfel T, Hohlfeld R, Meinl E. **2008**. Natalizumab disproportionately increases circulating pre-B and B cells in multiple sclerosis. Neurology 71:1350-1354.
150. von Andrian UH, Engelhardt B. **2003**. Alpha(4) integrins as therapeutic targets in autoimmune disease. N Eng J Med 348:68-72.
151. Gordon FH, Lai CWY, Hamilton MI, Allison MC, Srivastava ED, Fouweather MG, Donoghue S, Greenlees C, et al. **2001**. A randomized placebo-controlled trial of a humanized monoclonal antibody to alpha4 integrin in active Crohn's disease. Gastroenterology 121:268-274.
152. Miller DH, Khan OA, Sheremata WA, Blumhardt LD, Rice GPA, Libonati MA, Willmer-Hulme AJ, Dalton CM, et al. **2003**. A controlled trial of natalizumab for relapsing multiple sclerosis. N Engl J Med 348:15-23.
153. Bloomgren G, Richman S, Hotermans C, Subramanyam M, Goelz S, Natarajan A, Lee S, Plavina T, et al. **2012**. Risk of natalizumab-associated progressive multifocal leukoencephalopathy. N Engl J Med 366:1870-1880.
154. Ho PR, Koendgen H, Campbell N, Haddock B, Richman S, Chang I. **2017**. Risk of natalizumab-associated progressive multifocal leukoencephalopathy in patients with multiple sclerosis: A retrospective analysis of data from four clinical studies. Lancet Neurol 16:925-933.
155. Maloney DG, GrilloLopez AJ, White CA, Bodkin D, Schilder RJ, Neidhart JA, Janakiraman N, Foon KA, et al. **1997**. IDEC-C2B8 (Rituximab) anti-CD20 monoclonal antibody therapy in patients with relapsed low-grade non-Hodgkin's lymphoma. Blood 90:2188-2195.
156. Edwards JCW, Szczepanski L, Szechinski J, Filipowicz-Sosnowska A, Emery P, Close DR, Stevens RM, Shaw T. **2004**. Efficacy of B-cell-targeted therapy with rituximab in patients with rheumatoid arthritis. N Engl J Med 350:2572-2581.
157. Molloy ES, Calabrese LH. **2012**. Progressive multifocal leukoencephalopathy associated with immunosuppressive therapy in rheumatic diseases. Arthritis Rheum 64:3043-3051.

158. Hauser SL, Waubant E, Arnold DL, Vollmer T, Antel J, Fox RJ, Bar-Or A, Panzara M, et al. **2008.** B-cell depletion with rituximab in relapsing-remitting multiple sclerosis. N Engl J Med 358:676-688.
159. Benekli M, Hahn T, Shafi F, Qureshi A, Alam AR, Czuczman MS, Bernstein ZP, Chanan-Khan AA, et al. **2003**. Effect of rituximab on peripheral blood stem cell mobilization and engraftment kinetics in non-Hodgkin's lymphoma patients. Bone Marrow Transplant 32:139-143.
160. Keating MJ[1] CN, Messmer B, Damle RN, Allen SL, Rai KR, Ferrarini M, Kipps TJ. **2003**. Biology and treatment of chronic lymphocytic leukemia. Hematology Am Soc Hematol Educ Program 2003:153-75.
161. Coles AJ. **2013**. Alemtuzumab Therapy for Multiple Sclerosis. Neurotherapeutics 10:29-33.
162. Rowan W, Tite J, Topley P, Brett SJ. **1998**. Cross-linking of the CAMPATH-1 antigen (CD52) mediates growth inhibition in human B- and T-lymphoma cell lines, and subsequent emergence of CD52-deficient cells. Immunology 95:427-436.
163. D'Souza A, Wilson J, Mukherjee S, Jaiyesimi I. **2010**. Progressive multifocal leukoencephalopathy in chronic lymphocytic leukemia: A report of three cases and review of the literature. Clin Lymphoma Myeloma Leuk 10:E1-9.
164. Isidoro L, Rito L, Cordeiro G. **2014**. Progressive multifocal leukoencephalopathy in a patient with chronic lymphocytic leukaemia treated with alemtuzumab. BMJ Case Rep brc2013201781.
165. Kothary N, Diak IL, Brinker A, Bezabeh S, Avigan M, Dal Pan G. **2011**. Progressive multifocal leukoencephalopathy associated with efalizumab use in psoriasis patients. J Am Acad Dermatol 65:546-551.
166. Harel A, Horng S, Gustafson T, Ramineni A, Farber RS, Fabian M. **2018**. Successful treatment of progressive multifocal leukoencephalopathy with recombinant interleukin-7 and maraviroc in a patient with idiopathic CD4 lymphocytopenia. J Neurovirol 24:652-655.
167. Mullins C, Miranda J, Sandoval H, Ramos-Duran L, Tonarelli SB. **2018**. The benefit of mirtazapine in the treatment of progressive multifocal leukoencephalopathy in a young HIV-positive patient: A case report. Innov Clin Neurosci 15: 33–35.
168. Shelburne SA, Hamill RJ, Rodriguez-Barradas MC, Greenberg SB, Atmar RL, Musher DM, Gathe JC, Visnegarwala F, et al. **2002**. Immune reconstitution inflammatory syndrome: Emergence of a unique syndrome during highly active antiretroviral therapy. Medicine 81:213-227.
169. Tan K, Roda R, Ostrow L, McArthur J, Nath A. **2009**. PML-IRIS in patients with HIV infection: Clinical manifestations and treatment with steroids. Neurology 72:1458-1464.
170. Tan IL, McArthur JC, Clifford DB, Major EO, Nath A. **2011**. Immune reconstitution inflammatory syndrome in natalizumab-associated PML. Neurology 77:1061-1067.
171. Vaz B, Cinque P, Pickhardt M, Weber T. **2000**. Analysis of the transcriptional control region in progressive multifocal leukoencephalopathy. J Neurovirol 6:398-409.
172. Markowitz RB, Eaton BA, Kubik MF, Latorra D, McGregor JA, Dynan WS. **1991**. BK virus and JC virus shed during pregnancy have predominantly archetypal regulatory regions. J Virol 65:4515-4519.
173. Yogo Y, Kitamura T, Sugimoto C, Hara K, Iida T, Taguchi F, Tajima A, Kawabe K, Aso Y. **1991**. Sequence rearrangement in JC virus DNAs molecularly cloned from immunosuppressed renal-transplant patients. J Virol 65:2422-2428.
174. Agostini HT, Ryschkewitsch CF, Stoner GL. **1998**. Rearrangements of archetypal regulatory regions in JC virus genomes from urine. Res Virol 149:163-170.

175. Ault GS, Stoner GL. **1993**. Human polyomavirus JC promoter enhancer rearrangement patterns from proggresive multifocal leukoencephalopathy brain are unique derivatives of a single archetypal structure. J Gen Virol 74:1499-1507.
176. Reid CE, Li H, Sur G, Carmillo P, Bushnell S, Tizard R, McAuliffe M, Tonkin C, et al. **2011**. Sequencing and analysis of JC Virus DNA from natalizumab-treated PML patients. J Infect Dis 204:237-244.
177. Tan CS, Ellis LC, Wuthrich C, Ngo L, Broge TA, Saint-Aubyn J, Miller JS, Koralnik IJ. **2010**. JC virus latency in the brain and extraneural organs of patients with and without progressive multifocal leukoencephalopathy. J Virol 84:9200-9209.
178. Van Loy T, Thys K, Ryschkewitsch C, Lagatie O, Monaco MC, Major EO, Tritsmans L, Stuyver LJ. **2015**. JC Virus quasispecies analysis reveals a complex viral population underlying progressive multifocal leukoencephalopathy and supports viral dissemination via the hematogenous route. J Virol 89:1340-1347.
179. Marshall LJ, Dunham L, Major EO. **2010**. Transcription factor Spi-B binds unique sequences present in the tandem repeat promoter/enhancer of JC virus and supports viral activity. J Gen Virol 91:3042-3052.
180. Kumar KU, Pater A, Pater MM. **1993**. Human JC-virus perfect palindromic nuclear factor-1-binding sequences important for glial cell-specific expression in differentiating embryonal carcinoma-cells. J Virol 67:572-576.
181. Tan CS, Koralnik IJ. **2010**. Progressive multifocal leukoencephalopathy and other disorders caused by JC virus: Clinical features and pathogenesis. Lancet Neurol 9:425-437.
182. Gorelik L, Reid C, Testa M, Brickelmaier M, Bossolasco S, Pazzi A, Bestetti A, Carmillo P, et al **2011**. Progressive multifocal leukoencephalopathy (PML) development is associated with mutations in JC virus capsid protein VP1 that change its receptor specificity. J Infect Dis 204:103-114.
183. Takahashi K, Sekizuka T, Fukumoto H, Nakamichi K, Suzuki T, Sato Y, Hasegawa H, Kuroda M, et al. **2017**. Deep-sequence identification and role in virus replication of a JC virus quasispecies in patients with progressive multifocal leukoencephalopathy. J Virol 91:e01335-16.
184. Wharton KA, Quigley C, Themeles M, Dunstan RW, Doyle K, Cahir-McFarland E, Wei J, Buko A, et al. **2016**. JC polyomavirus abundance and distribution in progressive multifocal leukoencephalopathy (PML) brain tissue implicates myelin sheath in intracerebral dissemination of infection. Plos One 11:e0155897.
185. Nickeleit V, Hirsch HH, Binet IF, Gudat F, Prince O, Dalquen P, Thiel G, Mihatsch MJ. **1999**. Polyomavirus infection of renal allograft recipients: From latent infection to manifest disease. J Am Soc Nephrol 10:1080-1089.
186. Gard L, van Doesum W, Niesters HGM, van Son WJ, Diepstra A, Stegeman CA, Groen H, Riezebos-Brilman A, et al. **2017**. A delicate balance between rejection and BK polyomavirus associated nephropathy: A retrospective cohort study in renal transplant recipients. Plos One 12:e0178801.
187. Ramos E, Drachenberg CB, Papadimitriou JC, Hamze O, Fink JC, Klassen DK, Drachenberg RC, Wiland A, et al. **2002**. Clinical course of polyoma virus nephropathy in 67 renal transplant patients. J Am Soc Nephrol 13:2145-51.
188. Hirsch HH, Knowles W, Dickenmann M, Passweg J, Klimkait T, Mihatsch MJ, Steiger J. **2002**. Prospective study of polyomavirus type BK replication and nephropathy in renal-transplant recipients. N Engl J Med 347:488-496.
189. Hirsch HH, Brennan DC, Drachenberg CB, Ginevri F, Gordon J, Limaye AP, Mihatsch MJ, Nickeleit V, et al. **2005**. Polyomavirus-associated nephropathy

in renal transplantation: Interdisciplinary analyses and recommendations. Transplantation 79:1277-1286.
190. Viswesh V, Yost SE, Kaplan B. **2015**. The prevalence and implications of BK virus replication in non-renal solid organ transplant recipients: A systematic review. Transplantation Rev 29:175-180.
191. Sukov WR, Lewin M, Sethi S, Rakowski TA, Lager DJ. **2008**. BK virus-associated nephropathy in a patient with AIDS. Am J Kidney Dis 51:E15-18.
192. Bressollette-Bodin C, Coste-Burel M, Hourmant M, Sebille V, Andre-Garnier E, Imbert-Marcille BM. **2005**. A prospective longitudinal study of BK virus infection in 104 renal transplant recipients. Am J Transplant 5:1926-1933.
193. Brennan DC, Agha I, Bohl DL, Schnitzler MA, Hardinger HL, Lockwood M, Torrence S, Schuessler R, et al. **2005**. Incidence of BK with tacrolimus versus cyclosporine and impact of preemptive immunosuppression reduction. Am J Transplant 5:582-594.
194. Drachenberg CB, Papadimitriou JC, Hirsch HH, Wali R, Crowder C, Nogueira J, Cangro CB, Mendley S, Mian A, et al. **2004**. Histological patterns of polyomavirus nephropathy: Correlation with graft outcome and viral load. Am J Transplant 4:2082-2092.
195. Nickeleit V, Klimkait T, Binet IF, Dalquen P, Del Zenero V, Thiel G, Mihatsch MJ, Hirsch HH. **2000**. Testing for polyomavirus type BK DNA in plasma to identify renal-allograft recipients with viral nephropathy. N Engl J Med 342:1309-1315.
196. Nickeleit V, Singh HK, Randhawa P, Drachenberg CB, Bhatnagar R, Bracamonte E, Chang A, Chon WJ, et al. **2018**. The Banff working group classification of definitive polyomavirus nephropathy: Morphologic definitions and clinical correlations. J Am Soc Nephrol 29:680-693.
197. Kuypers DRJ, Vandooren AK, Lerut E, Evenepoel P, Claes K, Snoeck R, Naesens L, Vanrenterghem Y. **2005**. Adjuvant low-dose cidofovir therapy for BK polyomavirus interstitial nephritis in renal transplant recipients. Am J Transplant 5:1997-2004.
198. Sener A, House AA, Jevnikar AM, Boudville N, McAlister VC, Muirhead N, Rehman F, Luke PPW. **2006**. Intravenous immunoglobulin as a treatment for BK virus associated nephropathy: One-year follow-up of renal allograft recipients. Transplantation 81:117-120.
199. Wright AJ, Gill JS. **2016**. Strategies to prevent BK virus infection in kidney transplant recipients. Curr Opin Infect Dis 29:353-358.
200. Ramos E, Drachenberg CB, Wali R, Nogueira J, Cangro CB, Klassen DK, Papadimitriou JC. **2004**. Prospective biopsy for early diagnosis of polyomavirus-associated nephropathy (PVAN). Am J Transplant 4:198-198.
201. Sood P, Senanayake S, Sujeet K, Medipalli R, Van-Why SK, Cronin DC, Johnson CP, Hariharan S. **2013**. Donor and recipient BKV-specific IgG antibody and posttransplantation BKV infection: A prospective single-center study. Transplantation 95:896-902.
202. Bohl DL, Brennan DC, Ryschkewitsch C, Gaudreault-Keener M, Major EO, Storch GA. **2008**. BK virus antibody titers and intensity of infections after renal transplantation. J Clin Virol 43:184-189.
203. Rubinstein R, Pare N, Harley EH. **1987**. Structure and function of the transcriptional control region of nonpassaged BK virus. J Virol 61:1747-1750.
204. Moens U, Van Ghelue M. **2005**. Polymorphism in the genome of non-passaged human polyomavirus BK: Implications for cell tropism and the pathological role of the virus. Virology 331:209-231.
205. Sundsfjord A, Johansen T, Flaegstad T, Moens U, Villand P, Subramani S, Traavik T. **1990**. At least 2 types of control regions can be found among naturally-

occurring BK virus-strains. J Virol 64:3864-3871.
206. Sharma PM, Gupta G, Vats A, Shapiro R, Randhawa PS. **2007**. Polyomavirus BK non-coding control region rearrangements in health and disease. J Med Virol 79:1199-1207.
207. Helle F, Brochot E, Handala L, Martin E, Castelain S, Francois C, Duverlie G. **2017**. Biology of the BKPyV: An Update. Viruses 9:E327.
208. Peretti A, Geoghegan EM, Pastrana DV, Smola S, Feld P, Sauter M, Lohse S, Ramesh M, et al. **2018**. Characterization of BK polyomaviruses from kidney transplant recipients suggests a role for APOBEC3 in driving in-host virus evolution. Cell Host Microbe 23:628-35.
209. Krautkramer E, Klein TM, Sommerer C, Schnitzler P, Zeier M. **2009**. Mutations in the BC-Loop of the BKV VP1 region do not influence viral load in renal transplant patients. J Med Virol 81:75-81.
210. Barcena-Panero A, Van Ghelue M, Khan MTH, Echevarria JE, Fedele G, Moens U. **2012**. BK Virus-associated infection in cerebrospinal fluid of neurological patients and mutation analysis of the complete VP1 gene in different patient groups. J Cell Physiol 227:136-145.
211. Tremolada S, Delbue S, Castagnoli L, Allegrini S, Miglio U, Boldorini R, Elia F, Gordon J, Ferrante P. **2010**. Mutations in the external loops of BK virus VP1 and urine viral load in renal transplant recipients. J Cell Physiol 222:195-199.
212. Tremolada S, Akan S, Otte J, Khalili K, Ferrante P, Chaudhury PR, Woodle ES, Trofe-Clark J, et al. **2010**. Rare subtypes of BK virus are viable and frequently detected in renal transplant recipients with BK virus-associated nephropathy. Virology 404:312-318.
213. Boldorini R, Allegrini S, Miglio U, Paganotti A, Veggiani C, Mischitelli M, Monga G, Pietropaolo V. **2009**. Genomic mutations of viral protein 1 and BK virus nephropathy in kidney transplant recipients. J Med Virol 81:1385-1393.
214. Dugan AS, Gasparovic ML, Tsomaia N, Mierke DF, O'Hara BA, Manley K, Atwood WJ. **2007**. Identification of amino acid residues in BK virus VP1 that are critical for viability and growth. J Virol 81:11798-11808.
215. Cardoso KM, Diaz NC, Guimaraes M, Zalis MG, Delbue S, Ferrante P, Varella RB. **2015**. Genetic and structural analysis of polyomavirus BK T-antigens reveal a higher density of mutations at inter-domain and hexamerization regions, regardless the status of infection. J Med Virol 87:1418-1426.
216. Helanterä I, Hirsch HH, Auvinen E, Mannonen L, Nummi M, Wernli M, Ortiz F, Räisänen-Sokolowski A, et al. **2016**. High-level JCPyV viruria after kidney transplantation: Clinical and histopathological findings. J Clin Virol 85:75-79.
217. Janphram C, Worawichawong S, Disthabanchong S, Sumethkul V, Rotjanapan P. **2017**. Absence of JC polyomavirus (JCPyV) viremia in early post-transplant JCPyV nephropathy: A case report. Transplant Infect Dis 19:e12761.
218. Hirsch HH, Babel N, Comoli P, Friman V, Ginevri F, Jardine A, Lautenschlager I, Legendre C, et al. **2014**. European perspective on human polyomavirus infection, replication and disease in solid organ transplantation. Clin Microbiol Infect 20:74-88.
219. Lautenschlager I, Jahnukainen T, Kardas P, Lohi J, Auvinen E, Mannonen L, Dumoulin A, Hirsch HH, et al. **2014**. A case of primary JC polyomavirus infection-associated nephropathy. Am J Transplant 14:2887-2892.
220. Sanger F, Nicklen S, Coulson AR. **1977**. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A 74:5463-5467.
221. Hayden EC. **2014**. The $1,000 genome. Nature 507:294-295.
222. Koboldt DC, Larson DE, Sullivan LS, Bowne SJ, Steinberg KM, Churchill JD, Buhr AC, Nutter N, et al. **2014**. Exome-based mapping and variant prioritization for

inherited mendelian disorders. Am J Hum Gen 94:373-384.
223. Wheeler DA, Srinivasan M, Egholm M, Shen Y, Chen L, McGuire A, He W, Chen YJ, et al. **2008**. The complete genome of an individual by massively parallel DNA sequencing. Nature 452:872-U5.
224. Wang Z, Gerstein M, Snyder M. **2009**. RNA-Seq: A revolutionary tool for transcriptomics. Nature Rev Genet 10:57-63.
225. Laird PW. **2010**. Principles and challenges of genome-wide DNA methylation analysis. Nature Rev Genet 11:191-203.
226. Johnson DS, Mortazavi A, Myers RM, Wold B. **2007**. Genome-wide mapping of in vivo protein-DNA interactions. Science 316:1497-1502.
227. Macaulay IC, Voet T. **2014**. Single cell genomics: Advances and future perspectives. Plos Genetics 10:e1004126.
228. Minoche AE, Dohm JC, Himmelbauer H. **2011**. Evaluation of genomic high-throughput sequencing data generated on Illumina HiSeq and Genome Analyzer systems. Genome Biology 12:R112.
229. Merriman B, Rothberg JM, Ion Torrent R, Team D. **2012**. Progress in Ion Torrent semiconductor chip based sequencing. Electrophoresis 33:3397-3417.
230. Valouev A, Ichikawa J, Tonthat T, Stuart J, Ranade S, Peckham H, Zeng K, Malek JA, et al. **2008**. A high-resolution, nucleosome position map of C. elegans reveals a lack of universal sequence-dictated positioning. Genome Res 18:1051-1063.
231. Sims D, Sudbery I, Ilott NE, Heger A, Ponting CP. **2014**. Sequencing depth and coverage: Key considerations in genomic analyses. Nature Rev Genet 15:121-132.
232. Shendure J, Ji HL. **2008**. Next-generation DNA sequencing. Nature Biotechnol 26:1135-1145.
233. Quick J, Quinlan AR, Loman NJ. **2014**. A reference bacterial genome dataset generated on the MinION (TM) portable single-molecule nanopore sequencer. Gigascience 3:22.
234. Chaisson MJP, Huddleston J, Dennis MY, Sudmant PH, Malig M, Hormozdiari F, Antonacci F, Surti U, et al. **2015**. Resolving the complexity of the human genome using single-molecule sequencing. Nature 517:608-11.
235. Jain M, Koren S, Miga KH, Quick J, Rand AC, Sasani TA, Tyson JR, Beggs AD, et al. **2018**. Nanopore sequencing and assembly of a human genome with ultra-long reads. Nature Biotechnology 36:338-45.
236. Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, Peluso P, et al. **2009**. Real-time DNA sequencing from single polymerase molecules. Science 323:133-138.
237. Stoddart D, Heron AJ, Mikhailova E, Maglia G, Bayley H. **2009**. Single-nucleotide discrimination in immobilized DNA oligonucleotides with a biological nanopore. Proc Natl Acad Sci U S A 106:7702-7707.
238. Levene MJ, Korlach J, Turner SW, Foquet M, Craighead HG, Webb WW. **2003**. Zero-mode waveguides for single-molecule analysis at high concentrations. Science 299:682-686.
239. Travers KJ, Chin CS, Rank DR, Eid JS, Turner SW. **2010**. A flexible and efficient template format for circular consensus sequencing and SNP detection. Nucleic Acids Res 38:e159.
240. Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, et al. **2013**. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nature Methods 10:563-69.
241. Astier Y, Braha O, Bayley H. **2006**. Toward single molecule DNA sequencing: Direct identification of ribonucleoside and deoxyribonucleoside 5 '-monophosphates by using an engineered protein nanopore equipped with a molecular adapter.

J Am Chem Soc 128:1705-1710.
242. Jain M. **2017**. MinION Analysis and Reference Consortium: Phase 2 data release and analysis of R9.0 chemistry. F1000Res 6:760
243. Koren S, Schatz MC, Walenz BP, Martin J, Howard JT, Ganapathy G, Wang Z, Rasko DA, et al. **2012**. Hybrid error correction and de novo assembly of single-molecule sequencing reads. Nature Biotechnol 30:692-700.
244. Li H, Ruan J, Durbin R. **2008**. Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Res 18:1851-1858.
245. Trapnell C, Salzberg SL. **2009**. How to map billions of short reads onto genomes. Nature Biotechnol 27:455-457.
246. Broers SA. **2019**. Understanding and overcoming the pitfalls and biases of next-generation sequencing (NGS) methods for use in the routine clinical microbiological diagnostic laboratory. Eur J Clin Microbiol Infect Dis 38:1059-70.
247. Dohm JC, Lottaz C, Borodina T, Himmelbauer H. **2008**. Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. Nucleic Acids Res 36:e105.
248. Somasekar S, Lee D, Rule J, Naccache SN, Stone M, Busch MP, Sanders C, Lee WM, et al. **2017**. Viral surveillance in serum samples from patients with acute liver failure by metagenomic next-generation sequencing. Clin Infect Dis 65:1477-1485.
249. Dumoulin A, Hirsch HH. **2011**. Reevaluating and optimizing polyomavirus BK and JC real-time PCR assays to detect rare sequence polymorphisms. J Clin Microbiol 49:1382-1388.
250. Lundin S, Stranneheim H, Pettersson E, Klevebring D, Lundeberg J. **2010**. Increased throughput by parallelization of library preparation for massive sequencing. Plos One 5:e10029.
251. Livak KJ, Schmittgen TD. **2001**. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ method. Methods 25:402-408.
252. Henson JW. **1994**. Regulation of the glial-specific JC virus early promoter by the transcription factor-Sp1. J Biol Chemistry 269:1046-1050.
253. Kitamura T, Sugimoto C, Kato A, Ebihara H, Suzuki M, Taguchi F, Kawabe K, Yogo Y. **1997**. Persistent JC virus (JCV) infection is demonstrated by continuous shedding of the same JCV strains. J Clin Microbiol 35:1255-1257.
254. Kim J, Woolridge S, Biffi R, Borghi E, Lassak A, Ferrante P, Amini S, Khalili K, et al. **2003**. Members of the AP-1 family, c-Jun and c-Fos, functionally interact with JC virus early regulatory protein large T antigen. J Virol 77:5241-5252.
255. Wegner M, Drolet DW, Rosenfeld MG. **1993**. Regulation of JC virus by the pou-domain transcription factor TST-1: Implications for progressive multifocal leukoencephalopathy. Proc Natl Acad Sci U S A 90:4743-4747.
256. Monaco MCG, Sabath BF, Durham LC, Major EO. **2001**. JC virus multiplication in human hematopoietic progenitor cells requires the NF-1 class D transcription factor. J Virol 75:9687-9695.
257. Ravichandran V, Sabath BF, Jensen PN, Houff SA, Major EO. **2006**. Interactions between c-Jun, nuclear factor 1, and JC virus promoter sequences: Implications for viral tropism. J Virol 80:10506-10513.
258. Dyson N, Bernards R, Friend SH, Gooding LR, Hassell JA, Major EO, Pipas JM, Vandyke T, et al. **1990**. Large T-antigens of many polyomaviruses are able to form complexes with the retinoblastoma protein. J Virol 64:1353-1356.
259. Jordan JA, Manley K, Dugan AS, O'Hara BA, Atwood WJ. **2010**. Transcriptional regulation of BK Virus by nuclear factor of activated T cells. J Virol 84:1722-1730.

9 ORIGINAL PUBLICATIONS